Noise Induced Changes to Dynamic Behaviour of Stochastic Delay Differential Equations

Thesis submitted in accordance with the requirements of the University of Liverpool for the degree of Doctor in Philosophy by Stewart John Norton.

February 2008

Declaration

No part of the work referred to in this thesis has been submitted in support of an application for another degree or qualification of this or any other institution of learning. However some parts of the material contained herein have been previously published.

Acknowledgements

I would like to thank my colleagues in the Mathematics department at University College Chester for their support and encouragement during the period of the research presented in this thesis. In particular, I would like to acknowledge the advice and guidance of my colleague and supervisor Professor Neville J. Ford. Finally, to Moira, for her patience.

Abstract

This thesis is concerned with changes in the behaviour of solutions to parameter-dependent stochastic delay differential equations of the general form

$$(1)\quad Y(t) = Y(t_0) + \int_{t_0}^{t} F(s, Y(s), Y(s-\tau))ds + \int_{t_0}^{t} G(s, Y(s), Y(s-\tau))dW(s)$$

as the values of a parameter varies. This equation is often written, in the Itô sense, in the shorthand form

$$(2)\quad dY(t) = F(t, Y(t), Y(t-\tau))dt + G(t, Y(t), Y(t-\tau))dW(t), \qquad t \geq t_0$$

Since there is, in general, no closed-form solution available, one needs to consider numerical approximations. We consider the effect on these critical parameter values of

1. changing levels of noise,
2. the choice of approximation method,
3. the choice of step length.

Many of our results focus on the basic stochastic linear delay equation

$$\begin{aligned} dY(t) &= \lambda Y(t-\tau)dt + \mu Y(t)dW(t), \qquad t \geq 0 \\ (3)\qquad Y(t) &= \Phi(t), \qquad t \in [-\tau, 0]. \end{aligned}$$

and on the stochastic logistic delay equation

$$\begin{aligned} dY(t) &= \lambda Y(t-\tau)[1 + Y(t)]dt + \mu Y(t)dW(t), \qquad t \geq 0 \\ (4)\qquad Y(t) &= \Phi(t), \qquad t \in [-\tau, 0]. \end{aligned}$$

We conclude

1. a phenomenological approach to the investigation is insufficient for our purposes,
2. a dynamical approach based on estimates for Lyapunov exponents provides a more reliable underlying method,
3. the inter-relationship of critical parameter values, step length and choice of method can be established in a reliable way.

Finally, we give evidence that indicates that, as the level of noise increases, so the apparent order of a numerical method may be reduced.

Contents

Chapter 1

Introduction

Ordinary differential equations (ODEs) have long been used in many fields to model mathematically real life situations. Typically, examples can be found in modelling populations, (see [41, 16, 14]), diffusion [53], chemical kinetics [18], wave propagation [53], drug kinetics [11], biosciences [14], bending beams [33], electrical circuits [29], 'predator-prey' [18], planetary motion [18] and atomic waste disposal [14], to name just a few. As we try to improve and extend these models there are two natural paths that we can take. If we take the accepted logistic model [41] for populations we can take account of the gestation period of reproduction. To cope with this refinement we can add a delay to our ODE [28, 45, 70]

$$dP(t) = rP(t)\left[1 - \frac{P(t)}{E}\right]dt \tag{1.1}$$

to produce the delay differential equation (DDE)

$$dP(t) = rP(t)\left[1 - \frac{P(t-\tau)}{E}\right]dt \tag{1.2}$$

where τ is our delay or gestation period, E is the equilibrium population level and r is a constant depending upon proportionate birth and death rates. The deterministic equation needs an initial population value, P_0, but our delay equation needs an initial function defined on an interval of length τ. Most commonly, either $[-\tau, 0]$ or $[0, \tau]$ is used.

Our second path is to consider perturbations to our equations caused by random changes in the system. For instance we could just add a small random perturbation μdW, usually referred to by the term "noise"(see [73]), to equation (1.1) to give the equation

$$dP(t) = rP(t)\left[1 - \frac{P(t)}{E}\right]dt + \mu dW. \tag{1.3}$$

	Deterministic	Stochastic
Instantaneous response	Scalar ODEs Finite-dimensional; Solution depends on initial value;	Scalar SODEs Infinite-dimensional; Solution depends on initial value; Solution is a stochastic process;
Delayed response	Scalar DDEs Infinite-dimensional Solution depends on initial function;	Scalar SDDEs Infinite-dimensional; Solution depends on initial function; Solution is a stochastic process;

Table 1.1: Solution spaces of types of differential equations

The term dW is frequently taken as an increment of a Brownian motion path, and both dW and Brownian motion will be discussed fully in chapter 3. We note that in the equation (1.3) the noise term does not include the dependent variable P, and hence the equation is referred to as a stochastic differential equation (SODE) with additive noise. However, it may be more natural to consider our extension from equation (1.1) by looking at the proportionate population change $dP(t)/P(t)$ and adding our stochastic term to this quantity. This gives us

$$\frac{dP(t)}{P(t)} = r\left[1 - \frac{P(t)}{E}\right]dt$$

which now becomes

$$\frac{dP(t)}{P(t)} = r\left[1 - \frac{P(t)}{E}\right]dt + \mu dW.$$

Multiplying by $P(t)$ gives us the following SODE with multiplicative noise

$$dP(t) = rP(t)\left[1 - \frac{P(t)}{E}\right]dt + \mu P(t)dW. \tag{1.4}$$

This seems the more natural approach, and we will only consider equations with multiplicative noise in this thesis.

In this thesis we will be taking both of these paths together and consider stochastic delay differential equations (SDDEs). SDDEs can be approached either as SODEs with added delay or DDEs with noise [5]. We will take the second approach as work has already been published on bifurcations in DDEs. The table in [5] neatly sums up the solution spaces of the four types of differential equation and it is repeated here as table 1.1.

DDEs have infinite-dimensional solution space as a result of the infinite number of values of the initial functions and SODEs have infinite-dimensional solution space as a result of the infinite number of stochastic trajectories available. As

a consequence, a SDDE has both of these conditions, "increasing the scope of interesting bifurcation behaviour", [5]. It is bifurcations in SDDEs, and how they are perturbed by numerical methods of solution that is the main interest of this thesis.

1.1 ODEs and bifurcations

A general first order ODE is an equation of the form

$$y'(t) = f(t, y(t)), \qquad y(t_0) = y_0 \tag{1.5}$$

1.1.1 Existence and uniqueness of solutions

There are a number of theorems concerning the existence and uniqueness of such equations, see [11, 14, 22, 28, 76]. We quote a typical theorem, taken from [23], with conditions on the function $f(t, y(t))$:

Theorem 1.1.1 *Assume the following conditions are satisfied*

1. *The function $f(t, y)$ is continuous on the set D : $|t - t_0| \leq a$, $|y - y_o| \leq b$.*
2. *$f(t, y)$ satisfies the Lipschitz condition with respect to y*
   $|f(t, y_1) - f(t, y_2)| \leq L|y_1 - y_2|$.

Then, there exists a unique solution $y = y(t)$ of equation (1.5). This solution is defined on the interval

$$|t - t_0| \leq \delta = min\{a, \frac{b}{M}\}$$

where $M = sup|f(t, y)|$, $(t, y) \in D$.
Moreover, the solution is the uniform limit on $|t - t_0| \leq \delta$ of the sequence of functions $\{y_n(t)\}$, where $y_0(t) = y_0$ and

$$y_n(t) = y_0 + \int_{t_0}^{t} f(s, y_{n-1}(s))ds, \; n = 1, 2, \ldots$$

1.1.2 Stability

There are various descriptions of stability. As this thesis is not primarily dealing with ODEs we will only give some informal ideas at this stage, but see [11, 15, 23, 47] for full details. If we introduce small changes in the initial value and/or the parameters of the equation (1.5) then we say the solution is

1. *relatively stable* if the percentage change in the solution remains bounded,

2. *relatively asymptotically stable* if the percentage change in the solution tends to zero,

3. *stable* if the actual change in the solution remains bounded,

4. *asymptotically stable* if the actual change in the solution tends to zero.

1.1.3 Bifurcations in ODEs

The main aim of this thesis is to investigate the relationships between parameter values at changes in the behaviour of numerical solutions to stochastic delay differential equations. In particular we will look mostly at the linear equation, but we will make reference to the logistic equation. We will start in this chapter by introducing the concept of bifurcations, showing simple examples with the linear ODE. We will also demonstrate that the basic numerical ϑ-methods do not always have the same parameter values at bifurcations as predicted by the theory. Consider the linear ODE

$$y'(t) = \lambda y(t), \qquad y(t_0) = A \tag{1.6}$$

This equation has solution
$y(t) = Ae^{\lambda t}$
This solution bifurcates at $\Re(\lambda) = 0$ as

$$y(t) \longrightarrow \begin{cases} 0, & \Re(\lambda) < 0 \\ \text{oscillates}, & \Re(\lambda) = 0 \\ \text{is unbounded}, & \Re(\lambda) > 0 \end{cases}$$

For convergence to the stable solution we require a value of λ in the negative $\Re$ half plane. Figure 1.1 demonstrates this bifurcation at $\lambda = 0$.

1.2 Numerical methods

Numerical methods are considered fully in [60, 56, 47]. Equations of the form

$$y'(t) = f(t, y(t)) \qquad t \in [t_0, T] \tag{1.7}$$

will not, in general, have explicit solutions. We will more often than not need to find an approximation to the exact solution using a numerical method.
Brugnano and Trigiante (in [15]) state that we need the following three steps:

1. construct a mesh $\{t_i\}_{i=0}^N$ in the interval $[t_0, T]$,

2. replace the continuous problem by a discrete one, described on the discrete set (grid points) $\{t_i\}_{i=0}^N$,

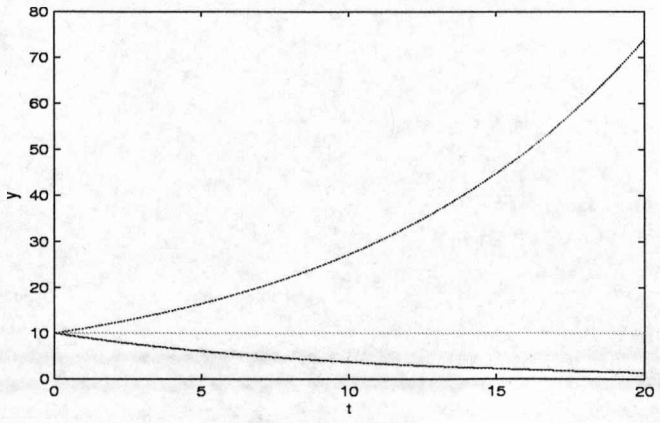


Figure 1.1: Bifurcation of the solution to $\frac{dy}{dt} = \lambda y$, $y(0) = 10$, at $\lambda = 0$.
Top curve $\lambda = 0.1$: Middle curve $\lambda = 0$: Bottom curve $\lambda = -0.1$.

3. solve the discrete problem.

The most simple grid points are $t_i = t_0 + ih, \qquad i = 1, 2, \ldots N$, where $h = \frac{T - t_0}{N}$ is the constant increment between consecutive grid points, called the step size.

"An easy to understand and implement method is the Euler method described by Euler in 1768", [47]. In this method we replace $y'(t)$ in equation (1.7) at $t = t_i$ by the simple approximation $\frac{y_{i+1} - y_i}{h}$ and hence we have

$$\frac{y_{i+1} - y_i}{h} = f(t_i, y(t_i)) = f(t_i, y_i).$$

Hence we have the recurrence formula

$$y_{i+1} = y_i + hf(t_i, y_i)$$

for $y(t)$ at all of the grid points. Iserles [56] states "the method can easily be extended to cater for variable step sizes as follows:

$$y_{i+1} = y_i + h_i f(t_i, y_i)\ "$$

How good is this approximation? We will discuss this question during this section, where we will also look at two types of numerical schemes, linear multistep formulae (LMF) and Runge-Kutta methods (RK).

1.2.1 Linear multistep formulae

Iserles [56] writes a general s-step method for equation (1.7) in the form

$$\sum_{m=0}^{s} a_m y_{i+m} = h \sum_{m=0}^{s} b_m f(t_{i+m}, y_{i+m}), \qquad i = 0, 1, \ldots, N \tag{1.8}$$

Iserles [56] also writes that the method can be characterised in terms of the polynomials

$$\rho(z) = \sum_{m=0}^{s} a_m z^m \qquad \sigma(z) = \sum_{m=0}^{s} b_m z^m \tag{1.9}$$

Clearly the Euler method classifies as a 1-step LMF since we can write it as

$$y_{i+1} - y_i = h[f(t_i, y_i)] = h[0.f(t_{i+1}, y_{i+1}) + 1.f(t_i, y_i)],$$

giving $\rho(z) = z - 1, \qquad \sigma(z) = 1$. Hairer et al [47] quote the implicit 2-step Adams-Moulton formula as:

$$y_{i+2} - y_{i+1} = h\left[\frac{5}{12}f(t_{i+2}, y_{i+2}) + \frac{2}{3}f(t_{i+1}, y_{i+1}) - \frac{1}{12}f(t_i, y_i)\right],$$

giving $\rho(z) = z^2 - z, \qquad \sigma(z) = \frac{5}{12}z^2 + \frac{2}{3}z - \frac{1}{12}$. We note that with this scheme we need to obtain a value for y_1 in addition to our initial value y_0 in order to initiate the recurrence for $y_2, \ldots, y_N$.

Brugnano and Trigiante [15] give the conditions for an LMF to be of order at least 1, (see the following definition on order). These are usually called the *consistency* conditions and can be written in the form

$$\rho(1) = 0, \qquad \rho'(1) = \sigma(1) \qquad\qquad (see\ [15]).$$

It is easy to check that both pairs of characteristic equations for the two methods given above satisfy the condition and so both have order of at least 1. We will see below that we can justify a higher order for the Adams-Moulton LMF.

Definition (2.3 in [47]).
An s-step LMF is said to be of *order* p if the local error, $y(t_i) - y_i$, is $\mathcal{O}(h^{p+1})$ for all sufficiently regular differential equations.

Theorem (2.4 in [47]).
An s-step LMF is of order p if and only if

1. $\sum_{m=0}^{s} a_m = 0$ and
2. $\sum_{m=0}^{s} a_m m^q = q \sum_{m=0}^{s} b_m m^{q-1}$ for $q = 1, 2, \ldots, p$

It is easy to verify that the coefficients of the Euler method in LMF format satisfy these conditions for $q = 1$, but not for $q = 2$, confirming the method is of order 1. We can also confirm that the implicit 2-step Adams-Moulton method's coefficients satisfy the conditions for $q = 1, 2, 3$ but not for $q = 4$, confirming that this method is of order 3. At this point we note that we will be referring to this Adams-Moulton LMF, and in particular its order, in chapter 4. Further work on this topic can be found in [18, 60, 72].

1.2.2 Runge-Kutta methods

To generate LMFs we used a numerical approximation for the derivative on the left hand side of equation (1.7). An alternative is to find a numerical approximation for the integral on the right hand side using quadrature formulae. For instance, if we integrate by the mid point method we get

$$y_{i+1} = y_i + hf(t_i + \frac{1}{2}h, \frac{1}{2}(y_{i+1} + y_i)).$$

"This gives a simple second order implicit mid point rule and a special case of the Runge-Kutta method", [56]. Runge-Kutta methods are described in some detail in [56, 47]. We will complete this section by listing the steps of a typical RK method. We will take Heun's order 3 explicit RK method listed in [47].

$$\begin{aligned} k_{i1} &= f(t_i, y_i), \\ k_{i2} &= f(t_i + \frac{1}{3}h, y_i + \frac{1}{3}k_{i1}), \\ k_{i3} &= f(t_i + \frac{2}{3}h, y_i + \frac{2}{3}k_{i2}), \\ y_{i+1} &= y_i + h(\frac{1}{4}k_{i1} + \frac{3}{4}k_{i3}). \end{aligned}$$

1.2.3 Summary

For this thesis we will be using the ϑ-methods, (see [56]). These methods are both 1-step LMF and RK methods [56, 47], and so make the obvious prototypes for our research.

1.2.4 ϑ-methods

The ϑ-methods applied to equation (1.5) produce the following recurrence equations

$$\begin{aligned} y_{n+1} &= y_n + (1-\vartheta)hf_n + \vartheta hf_{n+1} \\ y(0) &= y_0 \end{aligned}$$

For equation (1.6) these reduce to

$$\begin{aligned} y_{n+1} &= y_n + (1-\vartheta)h\lambda y_n + \vartheta h\lambda y_{n+1} \\ y(0) &= A, \end{aligned}$$

which can be written as

$$\begin{aligned} y_{n+1} &= \frac{[1+(1-\vartheta)h\lambda]}{[1-\vartheta h\lambda]} y_n \\ y(0) &= A, \end{aligned}$$

which finally gives us the numerical solution

$$y_n = \left[\frac{1+(1-\vartheta)h\lambda}{1-\vartheta h\lambda}\right]^n A. \tag{1.10}$$

We can now take the three methods with $\vartheta = 0$, 0.5 and 1, to investigate these solutions.

Question: Do the ϑ-methods preserve the bifurcation value of $\lambda = 0$ for the solutions to equation (1.6)?

1.2.5 Forward Euler, $\vartheta = 0$

The recurrence equation reduces to

$$y_n = [1 + h\lambda]^n A,$$

For stability we need $|1 + h\lambda| < 1$, so the region of stability needs $h\lambda$ to be contained in the unit circle, centre $z = -1$, of the complex plane. This is in contrast to the theoretic values, which require λ to be in the negative $\Re$ half plane. Figure 1.2 shows the difference between the theoretical and forward Euler stability regions. There is a large area not hatched in the negative half plane where the forward Euler method should, but does not, converge to the theoretical stable $y = 0$ solution.

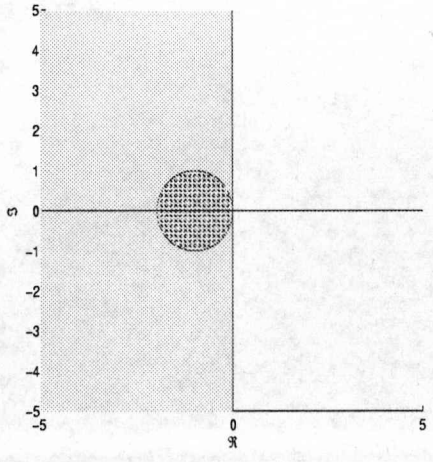


Figure 1.2: Stability region for $h\lambda$:
Theoretical, left hand half plane (shaded);
forward Euler, unit circle (cross hatch)

1.2.6 Backward Euler, $\vartheta = 1$

The recurrence equation reduces to

$$y_n = \frac{A}{[1 - h\lambda]^n},$$

For stability we need $|1 - h\lambda| > 1$, so the region of stability needs $h\lambda$ to be outside the unit circle, centre $z = 1$, of the complex plane. This is again in contrast to the theoretic values, as shown in figure 1.3, where there is a large area hatched in the real half plane showing the backward Euler method incorrectly converging to the stable solution of $y = 0$.

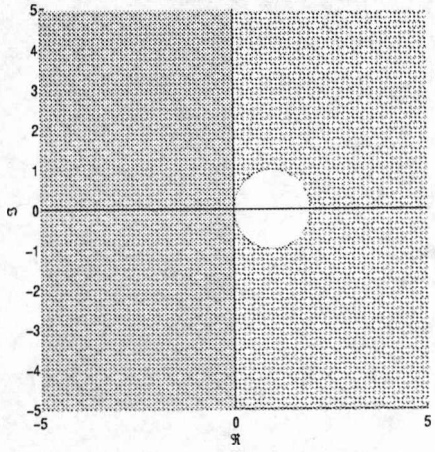


Figure 1.3: Stability region for $h\lambda$:
Theoretical, left hand half plane (darker shaded region);
backward Euler, outside unit circle (all shaded region)

1.2.7 Trapezium, $\vartheta = 0.5$

The recurrence equation reduces to

$$y_n = \left[\frac{1 + \frac{1}{2}h\lambda}{1 - \frac{1}{2}h\lambda}\right]^n A$$

For stability we need $\left|\frac{2+h\lambda}{2-h\lambda}\right| < 1$, so the region of stability needs $\Re(h\lambda) < 0$, which does correspond to the theoretical negative $\Re$ half plane, as $h > 0$.

1.2.8 Graphical comparisons

We have shown that even with the elementary equation (1.6) the basic numerical methods do not necessarily reproduce the theoretical solutions.

Forward Euler

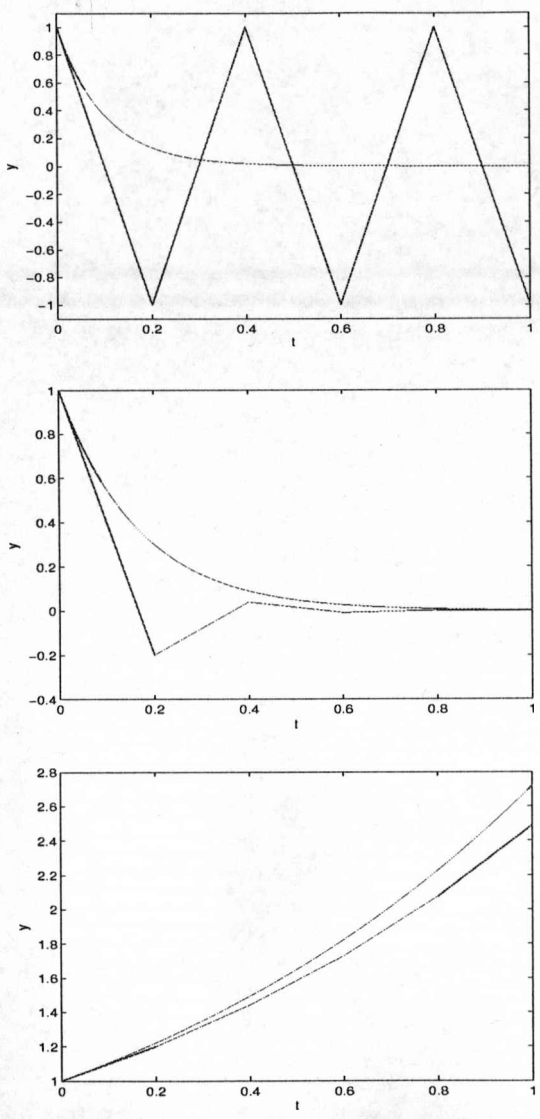


Figure 1.4: Theoretical and forward Euler solutions of $y'(t) = \lambda y(t), \quad y_0 = 1$
Top: $\lambda = -10,\ h = 0.2,\ h\lambda = -2$
Middle: $\lambda = -6,\ h = 0.2,\ h\lambda = -1.2$
Bottom: $\lambda = 1,\ h = 0.2,\ h\lambda = 0.2$

Figure 1.4 demonstrates the situation. The top figure has $h\lambda = -2$. This value is in the negative $\Re$ half plane so the theoretical solution has $y \to 0$. However, -2 is on the border of the unit circle, centre $z = -1$, and we can see that the forward Euler method is oscillating and not reflecting the true solution.

The middle figure has $h\lambda = -1.2$. This value is in the negative $\Re$ half plane and in unit circle, centre $z = -1$, so $y \to 0$ can be seen for both the theoretical and forward Euler graphs.

Finally, the bottom figure has $h\lambda = 1$. This value is outside of the stability regions for both the theoretical and forward Euler solutions, and consequently we can see $y \to \infty$ for both solutions.

Backward Euler

We refer to figure 1.5. The top graph has $h\lambda = -2$, which is in the negative $\Re$ plane, so the theoretical solution will converge to zero. The value is also outside the unit circle, centre $z = 1$, so the backward Euler method should also converge. The graph does in fact confirm that for both solutions we have $y \to 0$.

The middle graph has $h\lambda = 0.2$ which gives us a value in the positive $\Re$ plane and inside the relevant unit circle. As predicted by the theory in the previous section we can see that the solutions both show $y \to \infty$.

In the bottom diagram we have $h\lambda = 2.001$ in the positive $\Re$ plane, which theoretically should diverge. However, this value is outside the unit circle and the diagram confirms an incorrect convergence to zero.

Trapezium rule

Figure 1.6 confirms that the theoretical and numerical method diverge or converge together depending upon the sign of $h\lambda$.

1.2.9 Section summary

This section has demonstrated that the solution of the simple linear ODE undergoes a bifurcation as the parameter λ passes through zero, but that the basic ϑ-methods do not precisely replicate this phenomenon. We have included this section at some length to demonstrate this problem as the main theme of this thesis is to investigate the more complex situation with the linear stochastic delay differential equation.

1.3 DDEs and bifurcations

We will just consider equations where the delay is not state dependent. These are equations of the form

$$\begin{aligned} y'(t) &= f(t, y(t), y(t - \tau(t))) \qquad t \geq t_0, \\ y(t) &= \phi(t) \qquad t \leq t_0. \end{aligned} \tag{1.11}$$

There are many applications of DDEs, covering a wide range of models in population [24, 70], 'predator-prey' [45, 59], relativity [59], feedback control [45], nuclear

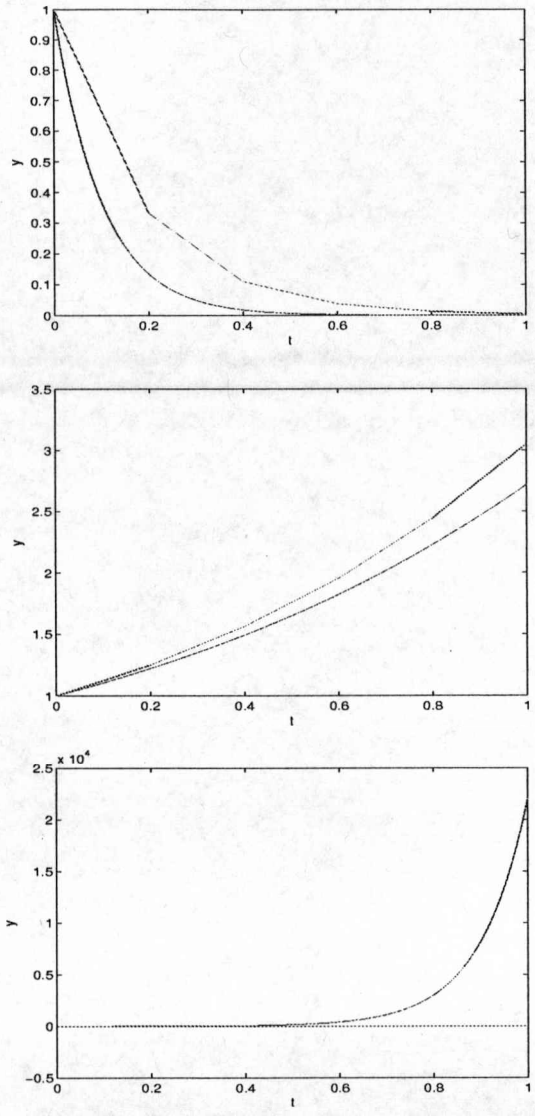


Figure 1.5: Theoretical and Backward Euler solutions of $y'(t) = \lambda y(t), \quad y_0 = 1$
Top: $\lambda = -10,\ h = 0.2,\ h\lambda = -2$
Middle: $\lambda = 1,\ h = 0.2,\ h\lambda = 0.2$
Bottom: $\lambda = 10.005,\ h = 0.2,\ h\lambda = 2.001$

reactors [59], biosciences [4], chemotherapy [59] and ecology [59], to name just a few.

1.3.1 Existence and uniqueness of solutions

We quote the theorem given in [12]

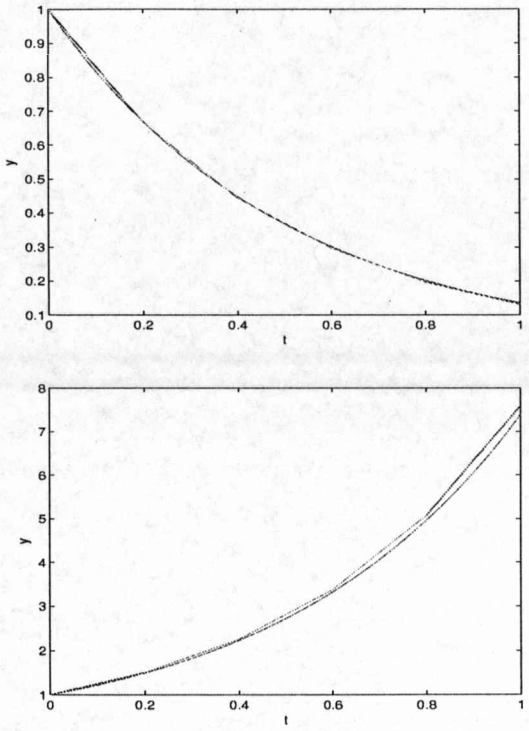


Figure 1.6: Theoretical and Trapezium rule solutions of $y'(t) = \lambda y(t), \quad y_0 = 1$
Top: $\lambda = -2,\ h = 0.2,\ h\lambda = -0.4$
Bottom: $\lambda = 2,\ h = 0.2,\ h\lambda = 0.4$

Theorem 1.3.1 *(Local existence) Consider the equation*

$$\begin{aligned} y'(t) &= f(t, y(t), y(t - \tau(t))) \qquad t \geq t_0, \\ y(t) &= \phi(t) \qquad t \leq t_0. \end{aligned}$$

Let $U \subseteq \Re^m$ and $V \subseteq \Re^m$ be neighbourhoods of $\phi(t_0)$ and $\phi(t_0 - \tau(t_0))$ respectively and assume that the function $f(t, u, v)$ is continuous with respect to t and Lipschitz continuous with respect to u and v in $[t_0, t_0 + h] \times U \times V$. Moreover, assume that the initial function $\phi(t)$ is Lipschitz continuous for $t \leq t_0$ and that the delay function $\tau(t) \geq 0$ is continuous with respect to t. Then the equation has a unique solution in $[t_0, t_0 + \delta)$ for some $\delta > 0$ and this solution continuously depends on the initial data.

Driver [28] proved this theorem with multiple delays. We refer to [12, 45, 48, 59] for more comprehensive analysis of the stability of DDEs.

1.4 Solutions of DDEs

In this section we will investigate methods for finding solutions to DDEs. In general, analytical solutions are not obtainable, so we will restrict our work to equations with constant delay, where it may be possible to use the method of steps to find solutions which can be used to compare numerical solutions. We will consider the equation

$$\begin{aligned} y'(t) &= f(t, y(t), y(t-\tau)) \qquad t \geq t_0, \\ y(t) &= \phi(t) \qquad t_0 - \tau \leq t \leq t_0 \end{aligned} \tag{1.12}$$

and also use the test equation

$$\begin{aligned} y'(t) &= -\frac{3}{2}y(t-1) \qquad t \geq 0, \\ y(t) &= t + \frac{1}{2} \qquad -1 \leq t \leq 0 \end{aligned} \tag{1.13}$$

We have chosen this test equation as it is a member of the family of deterministic equations that are examined in detail later in this thesis.

1.4.1 Method of steps

Details of this method can be found in [12, 28, 48]. We will apply the method to equation (1.12).
Step 1. Take $t \in [t_0, t_0 + \tau]$, so that $t - \tau \in [t_0 - \tau, t_0]$. We now have $y(t-\tau) = \phi(t-\tau)$. We can now substitute into the DDE to give

$$y'(t) = f(t, y(t), \phi(t-\tau))$$

which is an ODE. If $f(u, v, w)$ satisfies the conditions given earlier in theorem 1.1.1 then we can solve this ODE to give $y(t) = y_1(t)$ on $[t_0, t_0 + \tau]$.
Step 2. We now consider $t \in [t_0 + \tau, t_0 + 2\tau]$, so $t - \tau \in [t_0, t_0 + \tau]$, and hence $y(t-\tau) = y_1(t-\tau)$ and equation (1.12) becomes

$$y'(t) = f(t, y(t), y_1(t-\tau))$$

which is another ODE. Under the usual conditions we can solve this equation to give $y(t) = y_2(t)$ on $t \in [t_0 + \tau, t_0 + 2\tau]$.
Further steps. We now continue in this manner to find the solution $y(t) = y_n(t)$ for $t \in [t_0 + (n-1)\tau, t_0 + n\tau]$, where $y_n(t)$ is the solution of

$$y'(t) = f(t, y(t), y_{n-1}(t-\tau)).$$

Our complete solution is

$$y(t) = \begin{cases} y_1(t) & t \in [t_0, t_0 + \tau], \\ y_2(t) & t \in [t_0 + \tau, t_0 + 2\tau], \\ y_3(t) & t \in [t_0 + 2\tau, t_0 + 3\tau], \\ \dots & \\ y_n(t) & t \in [t_0 + (n-1)\tau, t_0 + n\tau], \\ \dots & \end{cases}$$

The solution propogates discontinuities at the ends of each interval. However, the solution can be shown to be *smoothing* as the number of steps increases. Usually, at t_0, $y(t)$ is continuous but $y'(t)$ is discontinuous. At $t_0 + \tau$, $y(t)$ and $y'(t)$ are continuous but $y''(t)$ is discontinuous. This smoothing continues so that at $t_0 + n\tau$, $y^{(n+1)}(t)$ is discontinuous while all of the lower derivatives are continuous, see [59, 12].

We can now demonstrate the technique by completing the first three steps with equation (1.13).
In the interval $[0, 1]$, the equation becomes

$$\begin{aligned} y'(t) &= -\frac{3}{2}y(t-1) \\ &= -\frac{3}{2}\phi(t-1) \\ &= -\frac{3}{2}(t-1) - \frac{3}{4} \\ y(0) &= \phi(0) \\ &= \frac{1}{2}. \end{aligned}$$

This can be solved to give

$$y_1(t) = -\frac{3}{4}(t-1)^2 - \frac{3}{4}t + \frac{5}{4}.$$

In the interval $[1, 2]$, the equation becomes

$$\begin{aligned} y'(t) &= -\frac{3}{2}y(t-1) \\ &= -\frac{3}{2}y_1(t-1) \\ &= \frac{9}{8}(t-2)^2 + \frac{9}{8}(t-1) - \frac{15}{8} \\ y(1) &= y_1(1) \\ &= \frac{1}{2}. \end{aligned}$$

This can be solved to give

$$y_2(t) = \frac{3}{8}(t-2)^3 + \frac{9}{16}(t-1)^2 - \frac{15}{8}t + \frac{11}{4}.$$

In the interval $[2,3]$, the equation becomes

$$\begin{aligned} y'(t) &= -\frac{3}{2}y(t-1) \\ &= -\frac{3}{2}y_2(t-1) \\ &= -\frac{9}{16}(t-3)^3 - \frac{27}{32}(t-2)^2 + \frac{45}{16}(t-1) - \frac{33}{8} \\ y(2) &= y_2(2) \\ &= -\frac{7}{16}. \end{aligned}$$

This can be solved to give

$$y_3(t) = -\frac{9}{64}(t-3)^4 - \frac{9}{32}(t-2)^3 + \frac{45}{32}(t-1)^2 - \frac{33}{8}t + \frac{419}{64}.$$

We can now check for discontinuities in our solutions.

At $t=0$, $\phi(0) = y_1(0) = \frac{1}{2}$,
but $\phi'(0) = 1 \neq y_1'(0) = \frac{3}{4}$. Hence there is a discontinuity at $y'(0)$.

At $t=1$, $y_1(1) = y_2(1) = \frac{1}{2}$, $\quad y_1'(1) = y_2'(1) = -\frac{3}{4}$,
but $y_1''(1) = -\frac{3}{2} \neq y_2''(1) = -\frac{9}{8}$. Hence there is a discontinuity at $y''(1)$.

At $t=2$, $y_2(2) = y_3(2) = -\frac{7}{6}$, $\quad y_2'(2) = y_3'(2) = -\frac{3}{4}$, $\quad y_2''(2) = y_3''(2) = \frac{9}{8}$,
but $y_2'''(2) = \frac{9}{4} \neq y_3'''(2) = \frac{27}{16}$. Hence there is a discontinuity at $y'''(2)$.

These results are exactly as predicted in our discussion about the smoothing of the solution.

1.4.2 Exponential solutions

For linear DDEs with constant delay we can look for solutions of the form $y(t) = Ae^{kt}$, see [11, 28, 24]. Consider the equation

$$\begin{aligned} y'(t) &= \psi y(t) + \lambda y(t-\tau), \qquad t \geq t_0 \\ y(t) &= \phi(t), \qquad t_0 - \tau \leq t \leq t_0 \end{aligned} \tag{1.14}$$

Substituting $y(t) = Ae^{kt}$ into equation (1.14) gives

$$Ake^{kt} = \psi Ae^{kt} + \lambda Ae^{k(t-\tau)}$$

which simplifies to

$$(k - \psi)e^{k\tau} - \lambda = 0. \tag{1.15}$$

This equation (1.15) is the *characteristic equation* and every root k_i of this equation gives a solution to equation (1.14) in the form $Ae^{k_i t}$. In fact, the general solution can be obtained this way as described in the following theorem.

Theorem 1.4.1 *(Taken from chapter 4 in [11]).*
Suppose that $\phi(t)$ is $C^1[t_0 - \tau, t_0]$. Let $\{k_i\}$ be a sequence of zeros of equation (1.15) arranged in order of decreasing real parts (or of increasing imaginary parts or absolute values). Then

$$y(t) = \sum_{i=1}^{\infty} p_i(t)e^{k_i t}, \qquad t \geq t_0 \tag{1.16}$$

is the solution of equation (1.14), where $p_i(t)$ is a polynomial of degree less than the multiplicity of the root $\{k_i\}$.

The expression $\sum_{i=1}^{\infty} e^{k_i t}p_i(t)$ is called the quasi-polynomial.
For our test equation (1.13) the characteristic equation is

$$ke^k + \frac{3}{2} = 0$$

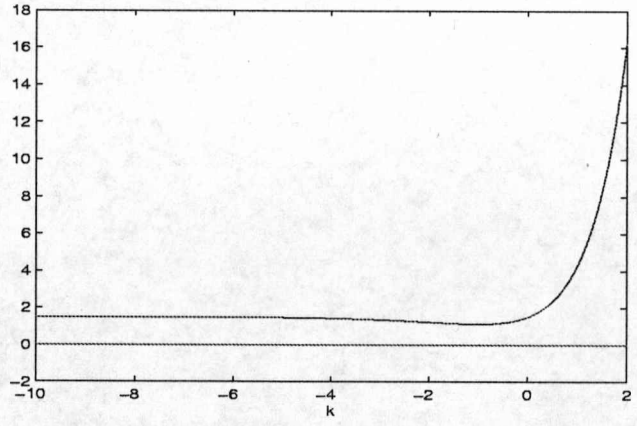


Figure 1.7: Graph showing the characteristic equation $ke^k + \frac{3}{2} = 0$ has no real roots

Figure 1.7 shows that there are no real roots to this equation. There are, however, infinitely many complex roots, so consequently this is not a practical method for solving our DDE. However, the approach gives useful information about solutions to DDEs.

1.4.3 Laplace transforms

Applying Laplace transforms to DDEs also gives us useful information about the solutions but again we do not usually obtain explicit solutions. We will apply the method on our test equation. We will write the Laplace transform $L(y)$ of $y(t)$ as $\bar{y}$, where

$$\bar{y} = \int_0^\infty y(t)e^{-st}dt$$

If we take Laplace transforms of $y'(t) = -\frac{3}{2}y(t-1)$ we get

$$\begin{aligned}
\int_0^\infty y'(t)e^{-st}dt &= \int_0^\infty -\frac{3}{2}y(t-1)e^{-st}dt \\
s\bar{y} - y(0) &= -\frac{3}{2}\int_{-1}^\infty y(u)e^{-s(u+1)}du \\
s\bar{y} - \frac{1}{2} &= -\frac{3}{2}e^{-s}\int_0^\infty y(u)e^{-su}du - \frac{3}{2}e^{-s}\int_{-1}^0 y(u)e^{-su}du \\
&= -\frac{3}{2}e^{-s}\bar{y} - \frac{3}{2}e^{-s}\int_{-1}^0 (u+\frac{1}{2})e^{-su}du \\
(s+\frac{3}{2}e^{-s})\bar{y} &= \frac{1}{2} - \frac{3}{2}e^{-s}[-\frac{1}{2s} - \frac{1}{s^2} - (\frac{1}{2s} - \frac{1}{s^2})e^s] \\
\therefore \quad \bar{y} &= \frac{\frac{1}{2} + \frac{3}{4s} - \frac{3}{2s^2} + (\frac{3}{4s} + \frac{3}{2s^2})e^{-s}}{s + \frac{3}{2}e^{-s}}
\end{aligned}$$

Hence we can find $y(t)$ as a contour integal as the inverse transform

$$y(t) = L^{-1}\left[\frac{\frac{1}{2} + \frac{3}{4s} - \frac{3}{2s^2} + (\frac{3}{4s} + \frac{3}{2s^2})e^{-s}}{s + \frac{3}{2}e^{-s}}\right].$$

Further work on Laplace transforms can be found in [11, 24, 29].

1.4.4 Numerical methods

We refer to section 4.1 in chapter 4 for the formula for several linear multistep methods for solving DDEs numerically. However, at this point it is interesting to compare the graphs of solutions using these methods, with different step lengths, with the exact solution using the method of steps.

Figure 1.8 shows that the second order trapezium rule and third order Adams Moulton rule coincide with the method of steps solution even with a step length of 0.1. All three solutions show as the green of the final solution added in the production of the graph. The two first order Euler methods improve considerably as h reduces from 0.1 to 0.01.

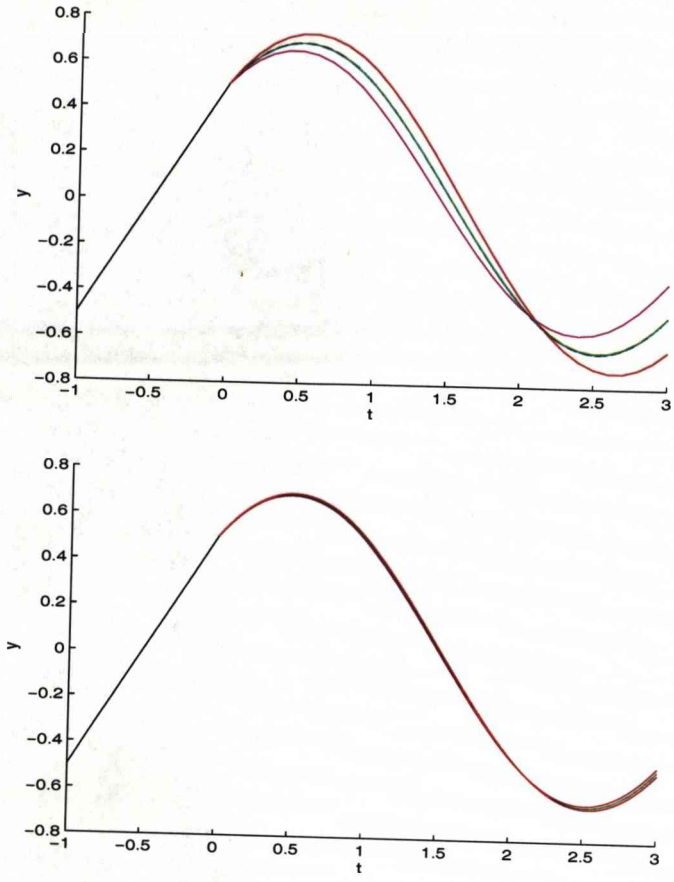


Figure 1.8: Graphs showing the convergence to the true solution as step length h decreases.
Top: $h = 0.1$ Bottom: $h = 0.01$
Exact (steps solution) - black; Adams Moulton - cyan; trapezium - green; backward Euler - magenta; forward Euler - red.

1.4.5 Computer packages

There are now quite a number of computer codes available for solving DDEs. *Archi*, [109], written by Paul in Fortran with the support of NAG, uses the Dormand and Prince fifth-order Runge-Kutta method. We used Archi in the application presented in the following chapter.

Bellen and Zennaro [12] gives a comprehensive list of software,which we repeat here with their references: CTMS [91, 92], DDE-STRIDE [83], DDVERK [85], DELH [93], DESOL [94], DIFSUB-DDE [82], DMRODE [88, 89], DRLAG6 [84], RADAR [86], RETARD [47] and SNDDELM [87].

DDE23 (Shampine & Thompson), [90], can be found as a code in MATLAB,

and is used as an example on our test equation. Figure 1.9 shows the result of running the basic DDE23 code supplied on MATLAB. Using the options facility to reduce the step length would improve the fit, especially at the turning points of the curve.

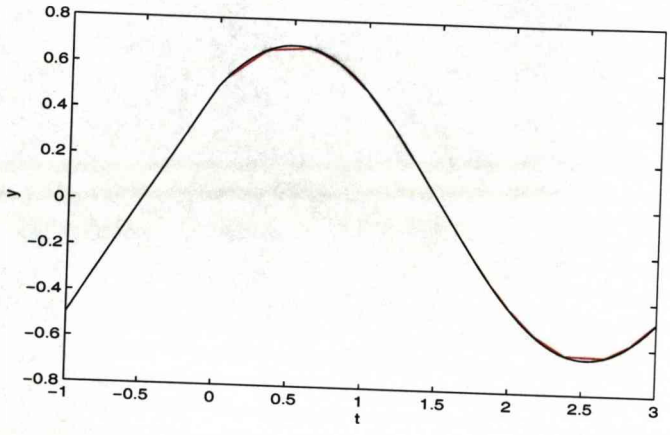


Figure 1.9: Basic DDE23 solution (red) compared with exact solution (black).

1.5 Stability

We will consider the criteria for stability in this section, restricting our review to equations with constant delay. We will use the equation

(1.17)
$$\begin{aligned} y'(t) &= f(t, y(t), y(t-\tau)) \qquad t \geq 0, \\ y(t) &= \phi(t) \qquad -\tau \leq t \leq 0 \end{aligned}$$

and define stability in the usual manner, quoting from [11]

Definition (Section 11.4 in [11]). Let x(t) be a function, continuous for $t > -\tau$, which satisfies the equation 1.17 for $t > 0$. This solution is said to be stable as $t \to \infty$ if, given two positive numbers t_0 and ϵ, there exists a corresponding positive number δ such that every continuous solution $y(t)$ of the equation 1.17 which satisfies

(1.18)
$$\max_{t_0 \leq t \leq t_0+\tau} |y(t) - x(t)| \leq \delta$$

will also satisfy

(1.19)
$$\max_{t_0 \leq t} |y(t) - x(t)| \leq \epsilon$$

The solution is said to be uniformly stable if, given ϵ, there exists a δ such that for any $t_0 \geq 0$ and any solution $y(t)$ which satisfies (1.18), $y(t)$ also satisfies (1.19).

Definition (Section 11.4 in [11]) The solution $x(t)$ above is said to be asymptotically stable if

1. it is stable;
2. for each $t_0 \geq 0$ there is a δ such that every solution $y(t)$ which satisfies (1.18) will also satisfy

$$\lim_{t \to \infty} |y(t) - x(t)| = 0. \tag{1.20}$$

It is said to be asymptotically stable in the large if it is stable and if every solution $y(t)$ satisfies the relation (1.20).

If we consider the linear DDE

$$y'(t) = \psi y(t) + \lambda y(t - \tau) \tag{1.21}$$

then we get the following result

Theorem 1.5.1 : *(Corollary 4.2 in [11]) A necessary and sufficient condition for all continuous solutions of equation 1.21 to approach zero as $t \to \infty$ is that all the roots of its characteristic equation, $e^{k\tau}(k - \psi) - \lambda = 0$, have negative real parts.*

The method of steps for DDEs with constant delay reduces the DDE to a set of ODEs over successive intervals of length τ. Consequently we can apply many of the results on uniqueness and stability of ODEs to this family of DDEs.

1.6 Bifurcations in DDEs

We showed in section 1.1 that the solutions to an ODE can bifurcate. We will show in this section that solutions of DDEs can also bifurcate at particular parameter values. If we consider the linear equation

$$\frac{dy(t)}{dt} = \lambda y(t - 1) \tag{1.22}$$

we see that the solutions diverge for $\lambda < -\frac{\pi}{2}$ and converge for $\lambda > -\frac{\pi}{2}$. Figure 1.10 clearly shows this phenomenon. The significance of the value $-\frac{\pi}{2}$ is emphasised in theorem 1.6.1.We also investigate this situation in detail in chapter 4.

Theorem 1.6.1 : *The steady state solution to the linear delay equation*

$$\frac{dY(t)}{dt} = -\lambda Y(t - \tau) \tag{1.23}$$

is stable if $0 \leq \lambda\tau \leq \frac{\pi}{2}$.

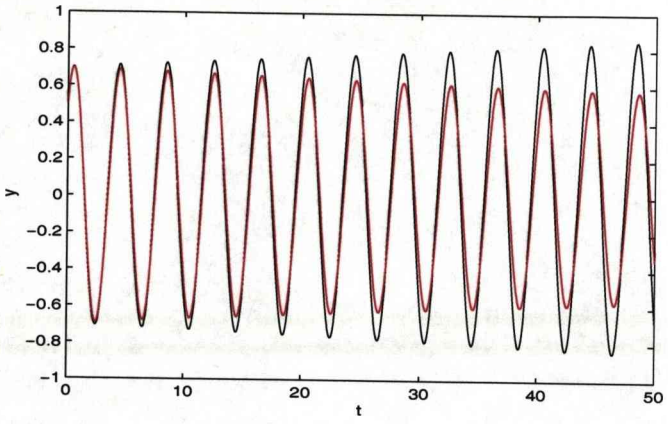


Figure 1.10: Solutions with $\lambda < -\frac{\pi}{2}$, diverging (black) and $\lambda > -\frac{\pi}{2}$, converging (red)

Proof: This result is central to this thesis so a proof, adapted from [70] is given in detail.

First, we non-dimensionalise the linear equation by writing $t^* = \lambda t$ and $\tau^* = \lambda\tau$. Substitution in our equation gives

$$\frac{dY}{dt^*} = -Y(t^* - \tau^*) \tag{1.24}$$

We look for solutions of the form $Y(t^*) = Ae^{kt^*}$, where A is a constant and, substituting in equation (1.24), the eigenvalues k are the roots of the equation

$$k = -e^{-k\tau^*} \tag{1.25}$$

The solutions of equation (1.25) are difficult to obtain. However, we are only interested in the solutions with $\Re(k) < 0$ which implies stability, as $Y(t^*)$ will decay exponentially in this case.
Now let $k = v + i\omega$ and take the modulus of (1.25) to get $|k| = e^{-v\tau^*}$. If $|k| \to \infty$ then $e^{-v\tau^*} \to \infty$, which requires $v \to -\infty$. Thus there must be a value v_0 such that $\Re(k) < v_0$.
We now introduce $z = \frac{1}{k}$ and define $f(z) = 1 + ze^{-\frac{\tau^*}{z}}$.
We note that $f(z)$ has an essential singularity at $z = 0$ and so, by Picard's theorem, $f(z) = 0$ has infinitely many complex roots in the neighbourhood of $z = 0$. As $f(z) = 0$ is equivalent to equation (1.25) we see that this equation has infinitely many roots k, with $\Re(k) < v_0$.
We now take the real and imaginary parts of equation (1.25) to give

$$v = -e^{-v\tau^*}cos(\omega\tau^*), \qquad \omega = e^{-v\tau^*}sin(\omega\tau^*) \tag{1.26}$$

and we require the range of values of τ^* such that $v < 0$. That is, we want to find the conditions such that the upper limit of v_0 on v is negative.
We take the two cases:
Case(i) k is real.
This requires $\omega = 0$, which clearly satisfies the second part of (1.26). Substituting into the first part of (1.26) gives us the equation $v = -e^{-v\tau^*}$. For $\tau^* > 0$ we can see graphically that there are no positive roots for v to this equation.
Case(ii): $\omega \neq 0$, so k is not real
Without loss of generality we can take $\omega > 0$, as if ω satisfies equations (1.26) then $-\omega$ is also a root.
Now the first part of equation (1.26) tells us that if $v < 0$ then $\omega < \frac{\pi}{2}$, since $-e^{-v\tau^*} < 0$ for all $v\tau^*$. We can now see that as τ^* increases from zero we first get $v = 0$ when $\omega\tau^* = \frac{\pi}{2}$. We can also see that in this case, and with $v = 0$, the only relevant solution to the second part of equation (1.26) is $\omega = 1$.
This now gives us the bifurcation situation where $\omega = 1$ and $\omega\tau^* = \tau^* = \frac{\pi}{2}$.
Hence we have the condition for stability as $0 < \tau^* < \frac{\pi}{2}$. returning to our original equation (1.23) we have $0 < \lambda\tau < \frac{\pi}{2}$.
Hence the condition for the steady state solution of equation (1.23) is

$$0 < \lambda < \frac{\pi}{2\tau}. \tag{1.27}$$

1.6.1 Numerical methods and bifurcations

In his PhD thesis [81] Wulf proved that when we apply a numerical scheme to solve equation (1.23) the position of the apparent bifurcation value of λ is perturbed from the theoretical value of $\frac{\pi}{2}$. Further, he showed that this perturbation is $\mathcal{O}(h^n)$, where h is the step length and n is the order of the numerical scheme. In chapter 4 we review this result and confirm it phenomenologically.

1.6.2 An example

We conclude this opening chapter by looking at the characteristic equation of

$$\frac{dy(t)}{dt} = \lambda y(t-1), \qquad \textit{see } [75] \tag{1.28}$$

Substituting $y(t) = Ae^{kt}$ we get

$$Ake^{kt} = \lambda Ae^{k(t-1)} \tag{1.29}$$

which leads to

$$k = \lambda e^{-k} \tag{1.30}$$

We have three situations in the case

1. $\lambda > 0$: In this case, equation (1.30) can be shown to have exactly one real root for k, giving a real solution for $y(t)$. However, equation (1.30) will have infinitely many complex roots and, as we will see below, some of these will lead to real solutions $y(t)$. As stated in section 1.4.2 the general solution is made up of a combination (depending upon the initial function $\phi(t)$) of these solutions.

2. $\lambda = 0$: The solution is the trivial solution, $k = 0$, corresponding to $y(t) = A$.

3. $\lambda < 0$: There are three cases

   (a) $\lambda \in (-e^{-1}, 0)$: There are two real solutions k_1, k_2 of equation (1.30), satisfying $k_1 < -1 < k_2 < 0$, together with infinitely many complex roots.

   (b) $\lambda = -e^{-1}$: We have exact one real root, $k = -1$, together with the complex roots. We note that for this value of λ if the initial function is $\phi(t) = Ae^{-t}$ then the solution of equation (1.28) can be expressed explicitly as $y(t) = Ae^{-t}$.

   (c) $\lambda < -e^{-1}$: For this final case it is easy to show that we have no real solutions to equation (1.30). We only have the infinite number of complex solutions. However, as we will show below, this does not mean that there are no real solutions to our DDE.
   Take $\lambda = -\frac{\pi}{2} < -e^{-1}$, which is our critical value in the bifurcation theory. $k = \pm\frac{\pi}{2}i$ are solutions of the characteristic equation. Hence we have
   $$y(t) = A\cos(\frac{\pi}{2}t) + B\sin(\frac{\pi}{2}t)$$
   as a real solution of
   $$\frac{dy(t)}{dt} = -\frac{\pi}{2}y(t-1).$$

Chapter 2

An application

In our paper [3] we looked at the progression from ODEs to DDEs in immunology model selection, and we used computational approaches to estimate the parameters to establish the best choice of model. Relevant sections of our paper are included here to demonstrate the technique:

"Abstract: One of the significant challenges in biomathematics (and other areas of science) is to formulate meaningful mathematical models. Our problem is to decide on a parametrized model which is, in some sense, most likely to represent the information in a set of observed data. In this paper, we illustrate the computational implementation of an information-theoretic approach (associated with a maximum likelihood treatment) to modelling in immunology.

The approach is illustrated by modelling LCMV infection using a family of models based on systems of ordinary differential and delay differential equations. The models (which use parameters that have a scientific interpretation) are chosen to fit data arising from experimental studies of virus-cytotoxic T lymphocyte kinetics; the parametrized models that result are arranged in a hierarchy by the computation of Akaike indices. The practical illustration is used to convey more general insight. Because the mathematical equations that comprise the models are solved numerically, the accuracy in the computation has a bearing on the outcome, and we address this and other practical details in our discussion."

The abstract refers to Akaike indices for finding the hierarchy of the models, and we define these indices in section 2.1 below. We follow on with further sections of the paper, to complete this chapter.

2.1 Sections from Immunology paper [3]

The paper is reproduced from Section 2.0.3

"A scoring mechanism: Given a family of models (each one with a best-fit set of parameters), the question is how to rank them by giving each a score (thus arriving at a hierarchy of parametrized models). The goodness of fit associated

with parameter estimates $\widetilde{p}$ can be characterized, when one has confidence in the form of the model, by the size of an objective function $\Phi_\star(\widetilde{p})$. This is the data-fitting approach, and here $\widetilde{p}$ may be an approximation (however obtained) to $\widehat{p}$ such that $\Phi_\star(\widehat{p}) = \min_p \Phi_\star(p)$. Thus, one criterion by which to judge a model may be the size of $\Phi_\star(\widetilde{p})$ (see [100]). However, if there is a *number* of candidate models, our task is not simply to identify one with the smallest objective function but to incorporate other criteria for discriminating between models of differing complexity. There are (information-theoretic) criteria, such as the Akaike, Schwarz, and Takeuchi information criteria[1] and generalizations related to informational complexity of models, which depend not only upon the maximum likelihood estimation bias [95, 102, 110] but incorporate the *number of parameters* and the *number of observations* in a quantitative evaluation of different models. Burnham and Anderson [102] review, as a natural basis for model selection, both the concept of K-L information and maximum likelihood.

For the Akaike and the corrected Akaike criteria, the indicators are the size of the measures μ_{AIC} and μ_{cAIC} given by

$$\text{(2.1a)} \quad \mu_{AIC} = -2 \ln \mathcal{L}(\widehat{p}) + 2(L+1),$$

$$\text{(2.1b)} \quad \mu_{cAIC} = -2 \ln \mathcal{L}(\widehat{p}) + 2(L+1) + \frac{2(L+1)(L+2)}{n-L-2}, \text{ with } n = NM,$$

respectively; see [102]. These indicators are expressed in terms of the MLE $\mathcal{L}(\widehat{p})$. There are $L+1$ parameters being estimated, comprising $p_1, p_2, \ldots, p_L$ and σ, since we currently assume that a single value σ, which we also estimate, characterizes all the variances. The advice quoted by Burnham and Anderson in [102] is that (2.1a) is satisfactory if $n > 40(L+1)$, otherwise (2.1b) is preferred by these authors. As $n \to \infty$, $\mu_{cAIC} \to \mu_{AIC}$.

Our interest is in the relative size of the indicators; thus (omitting technical details) it is convenient to discard extraneous terms and employ the revised indicators

$$\text{(2.2a)} \qquad \breve{\mu}_{AIC} = n \cdot \ln(\Phi_{\Omega LS}(\widehat{p})) + 2(L+1),$$

$$\text{(2.2b)} \qquad \breve{\mu}_{cAIC} = \breve{\mu}_{AIC} + \frac{2(L+1)(L+2)}{n-L-2}$$

where $\ell n(\ )$ denotes the natural logarithm,

(2.2c)

$n = NM$ (the sample size), and L is the number of parameters estimated.

[1]The Akaike criterion is based upon the Kullback-Leibler notion of information or distance between two probabilistic models (information loss) [107] approximated using the maximum likelihood estimation [95, 102]

2.2 Experimental LCMV Infection

2.2.1 Immune response to infection

The infection of a mouse with Lymphocytic Choriomeningitis Virus (LCMV) provides a basic experimental system used in immunology to address fundamental issues of virus-host interaction [111]. The infection results in the activation of immune responses and clonal burst [101] of virus-specific cytotoxic T-lymphocytes (CTL). We note that Ehl et al. [105] observed that "The use of a well-characterized murine infectious disease, which has been shown to be almost exclusively controlled by CTL-mediated perforin-dependent cytotoxicity, provides an exceptionally solid basis for the formulation of [models]".

At discrete times, it is possible to measure, experimentally, (*i*) the amount of the virus, measured in plaque forming units (*pfu*), and (*ii*) the virus-specific CTL (measured in the number of cells found per spleen[2]).

2.2.2 The experimental framework and the observations

In general, it is possible that data comes from a single experiment, or that the data arises from several experiments or a series of observations. Our mathematical models rely upon data being of a certain type: we assume the mean values of data are, at each time, normally or log-normally distributed, and independent.

The experimental data is provided in table 2.1. It was obtained as follows. A batch of genetically identical C57BL/6 mice were infected with 200 *pfu* (plaque forming units) of LCMV (WE strain), delivered intravenously. Viral titers in spleens were determined at days 1, 2, 3, 4, 6, 8, 10, 12 and 14 days post-infection and the clonal expansion of CTL cells specific for the gp33 epitope in spleens was assessed using tetramer analysis (see below). The techniques are standard; see for example [96, 99]. At the indicated time-points after infection, two mice were bled and single cell suspensions were prepared of spleen, prior to the determination of absolute cell counts using FACS and Neubauer equipment.

An important feature is that the mice were genetically identical, produced by inbreeding. Inbred strains reduce experimental variation; their immune responses can be studied in the absence of variables introduced by individual genetic difference. When the mice are genetically identical, it is argued that large numbers of mice are not required and the mean obtained represents the mean of a larger set of data. This assertion merits closer examination and testing, but we proceed on the basis that it is correct.

For reliable parameter estimation it is useful to have an idea of the CTL kinetics at times earlier than 6 days post infection – before the virus population starts to decrease. The quantity of virus-specific CTL below 5000 cell/spleen

[2]Some modellers introduce as a variable the amount of virus-specific memory CTL, a subset of (*ii*) that is harder to quantify reliably.

	$V(t)$		$E(t)$	
Time (days)	Set 1 - Virus Population (*pfu*)	Set 2 - Virus Population (*pfu*)	Set 1 CTLs Population (cells)	Set 2 - CTL Population (cells)
1	3.55×10^4	1.20×10^4	b.d.l.	b.d.l.
2	5.0×10^5	1.6×10^6	b.d.l.	b.d.l.
3	3.8×10^6	3.9×10^6	b.d.l.	b.d.l.
4	3.2×10^6	2.1×10^6	b.d.l.	b.d.l.
6	3.7×10^4	1.25×10^5	8.33×10^5	9.85×10^5
8	3.1×10^4	2.6×10^4	4.75×10^6	4.03×10^6
10	2.5×10^2	8.0×10^4	4.16×10^6	5.8×10^6
12	2.0×10^2	7.5×10^2	3.07×10^6	2.25×10^6
14	b.d.l.	b.d.l.	2.22×10^6	2.89×10^6

"b.d.l." *means "below the detection limit".*

Table 2.1: Data set for the virus and cytotoxic T lymphocyte kinetics in the spleen after systemic infection with 200 *pfu* of LCMV-WE

cannot be detected using the tetramer technique. Our experience (arising from numerous studies with the LCMV system) suggests that after injection of 200 *pfu* of LCMV the proliferating CTLs should reach the detection threshold at about two and a half days. This evidence was considered in the parameter estimation, by supplementing table 2.1 with a CTL reading (representing the least possible detection level) at day 2.5.

The detection threshold for LCMV in the spleen is about 100 *pfu*. LCMV-WE dropped below the detection threshold by day 14; however, it is believed that the virus still persisted below the detection level for some time. To ensure that the LCMV number in the model remains below the detection threshold, between days 12 to 14 we supplement the data with an assumption that the virus quantity on day 14 was 10 *pfu*/spleen.

2.3 Hierarchy of mathematical models

A priori immunological and mathematical knowledge enter the models in the form of simplifying assumptions. Potentially, the interaction between virus and immune system can be described by multiple mechanisms and considering various sets of differential equations. One may thus argue that different mechanisms and their functional forms might equally well describe the data set and the goodness-of-fit (*i.e.*, the maximized likelihood function) is not sufficient to judge whether the model is correct. It has been observed elsewhere that the maximum likelihood principle leads to choosing the models with higher possible complexity

(corresponding to more parameters) [110]. If there is a number of candidate models, our task is not simply to identify the one with the smallest objective function but to consider the principle of parsimony in model evaluation, and the maximum use of information implicit in the data. We suggest a hierarchy of mathematical models that were distilled from the existing literature.

The mathematical models for the virus-CTL interaction in LCMV infection are defined within a set of two or three dimensional ODEs or DDEs for the evolution of the virus, *V(t)*, and virus-specific CTL (activated and memory cells - $E(t), E_m(t)$) population dynamics.

The equation for the rate of change of the virus population is the same for all the models and is based upon a Verhulst-Pearl logistic growth term and second order elimination kinetics. The models differ in the way the immune response is described - an issue of some controversy in today's mathematical immunology. Specifically, the models differ with respect to the following building blocks:

1. virus-dependent CTL proliferation (basic predator-prey versus the Holling type II response);
2. whether a time-lag in the division of CTL (cell division time) is included;
3. consideration of homeostasis for naive CTL precursors;
4. whether a separate equation for the memory CTL is used.

The death rate of CTL is assumed constant. Overall we consider the following five models that have their counterparts in the literature. See table 2.2 for biological definitions of the parameters included in the model.

Model 1 (M_1): (simplest predator-prey consideration of the CTL dynamics)

$$\frac{d}{dt}V(t) = \beta \cdot V(t) \cdot \left(1 - \frac{V(t)}{K}\right) - \gamma \cdot V(t) \cdot E(t) \tag{2.3}$$

$$\frac{d}{dt}E(t) = b_1 \cdot V(t) \cdot E(t) - \alpha_E \cdot E(t) \tag{2.4}$$

Model 2 (M_2): (virus-dependent with saturation CTL proliferation)

$$\frac{d}{dt}V(t) = \beta \cdot V(t) \cdot \left(1 - \frac{V(t)}{K}\right) - \gamma \cdot V(t) \cdot E(t) \tag{2.5}$$

$$\frac{d}{dt}E(t) = \underbrace{b_2 \cdot V(t) \cdot E(t)/(\vartheta_{Sat} + V(t))}_{\text{A modification of model 1}} - \alpha_E \cdot E(t) \tag{2.6}$$

Model 3 (M_3): (virus-dependent with saturation CTL proliferation with time lag)

$$\frac{d}{dt}V(t) = \beta \cdot V(t) \cdot \left(1 - \frac{V(t)}{K}\right) - \gamma \cdot V(t) \cdot E(t) \tag{2.7}$$

$$\frac{d}{dt}E(t) = \underbrace{b_3 \cdot V(t-\tau) \cdot E(t-\tau)/(\vartheta_{Sat} + V(t))}_{\text{As in model 2 but incorporating delay}} - \alpha_E \cdot E(t) \tag{2.8}$$

Model 4 (M_4): (primary CTL homeostasis)

$$\frac{d}{dt}V(t) = \beta \cdot V(t) \cdot \left(1 - \frac{V(t)}{K}\right) - \gamma \cdot V(t) \cdot E(t) \tag{2.9}$$

$$\frac{d}{dt}E(t) = b_4 \cdot V(t-\tau) \cdot E(t-\tau)/(\vartheta_{Sat} + V(t)) - \underbrace{\alpha_E \cdot E(t) + T^*}_{\text{includes additive term}} \tag{2.10}$$

Model 5 (M_5): (Additional equation for the population of memory CTL)

$$\frac{d}{dt}V(t) = \beta \cdot V(t) \cdot \left(1 - \frac{V(t)}{K}\right) - \gamma \cdot V(t) \cdot E(t) \tag{2.11}$$

$$\frac{d}{dt}E(t) = b_5 \cdot V(t-\tau) \cdot E(t-\tau)/(\vartheta_{Sat} + V(t)) - \alpha_E \cdot E(t) - r_m \cdot E(t) + T^* \tag{2.12}$$

$$\frac{d}{dt}E_m(t) = r_m \cdot E(t) - \alpha_m \cdot E_m(t) \tag{2.13}$$

Observe that Model 1 is not a special case of Model 2 (the models are not nested); however, Models 2 to 5 are nested.

Remark 2.3.1 The parameters in all the above models are meaningful only if they have non-negative values. This can be accommodated by, for example, writing $p_\ell^{\star 2}$ for every non-negative parameter p_ℓ, where $p_\ell^\star = \sqrt{p}_\ell$, in an unconstrained optimization. (The parameters to be recovered remain the original $\{p_\ell\}$.)

In this study, we take as initial data

$$V(t) = 0,\ t \in [-\tau, 0),\ V(0) = V_0;\ E(t) = E_0,\ t \in [-\tau, 0];\ E_m(0) = 0;$$

(with $t_0 = 0$) and

$$V_0 = 200 \text{ and } E_0 = 265.$$

Here, V_0 and E_0 are the initial values for the dose of infection (measured in pfu) and the number of naive CTL (measured in cells). These parameters were considered to be fixed. The problem of identifying initial data is not addressed here; see [98], the references therein, and related work.

Parameter (units) The units are d (days), *pfu* (plaque forming units)	Notation
Virus exponential growth rate (d^{-1})	β
Carrying capacity for the virus (copies/spleen)	K
Virus elimination rate (1/copy/d)	γ
CTL stimulation rate (1/copy/d, M_1; d^{-1}, M_2 to M_5)	b_i
CTL division time (d)	τ
Viral load for half-maximal CTL stimulation (copy/spleen)	ϑ_{Sat}
Death rate of CTL (d^{-1})	α_E
Specific precursor CTL export from thymus (cell/spleen/d)	T^*
Reversion activated CTL into the memory state (d^{-1})	r_m
Death rate of memory CTL (d^{-1})	α_m

Table 2.2: Biological definition of the model parameters for virus-CTL dynamics in the spleen during primary LCMV infection. The spleen volume is estimated to be about 0.1 *(milliliters)*.

2.4 Computational methodology and numerical results

2.4.1 Description of numerical techniques

Minimization of a functional dependent on the solution of a system of ODEs or DDEs is based on two components: (*i*) an ability to find the numerical solution of the DDE (and thereby evaluate the objective function) and (*ii*) an ability to find the parameters that provide a global minimum of an objective function when there are constraints on the parameters (eg., $\tau \geq 0$). The availability of codes, and their suitability, depends upon the computer system (eg., Windows or Linux) and the computer package or language (eg., MATLAB or a version of FORTRAN).

In similar works, a degree of "sanitization" sometimes occurs: it is rarely mentioned that the objective function that is to be minimized is not computed exactly[3] (in our case, the solution of an ODE or DDE is computed using certain prescribed tolerances that govern the accuracy); see §2.4.2. In the illustrations below, we shall present differing results arising from different tolerances.

The user of a "black-box" code generally has the task of selecting tolerances that govern the accuracy of the computation. In practice, it is desirable to choose a tolerance sufficiently small that one has confidence in the parameters (it is

[3]It has to be remarked that the determination of error estimates, or confidence limits, can be as computationally expensive as the calculation of the approximation (possibly more expensive).

possible to compute confidence intervals), but large enough that the calculations are tractable given the demand on computer memory and processing speed. In a well-written code, tolerances are input via a parameter list. In some codes, it could prove necessary to "take the cover off the black box" (thus invalidating any warranty from the author!!) in order to exercise the required options.

2.4.2 Numerical methods for ODEs and DDEs

There is a wide choice of codes available for the numerical solution of ODEs. The situation for DDEs is less satisfactory. Most (if not all) codes for evolutionary problems proceed in an evolutionary mode, computing the solution over successive steps $[t_n, t_{n+1}]$ using a "step" of length $h_n = t_{n+1} - t_n$. For DDE codes we refer to, eg.,

- `http://www.mathworks.com/access/helpdesk/help/techdoc/ref/dde23.html` for MATLAB,
- the code `Archi`, at `www.ma.man.ac.uk/~chris/reports/rep283.pdf`, for FORTRAN (and the related codes `Archi-L`, `Archi-N`) and
- the code `RETARD` (the initial version to be found in [106]).

The parameters in ODE and DDE solvers usually include an "error-per-step" tolerance (`eps`). Codes for ODEs are usually optimized for "non-stiff" problems or "stiff" problems; a few are type-insensitive. Robust DDE codes that can cope with behaviour akin to "stiffness" are, at the time of writing, wanting. A practical difficulty when solving DDEs arises from the need to store the "history" (an approximation over an interval of the form $[t - \tau, t]$) which imposes large storage demands when the step-size is small relative to τ.

2.4.3 Numerical methods for optimization

Computer codes available for optimization exist in MATLAB and in FORTRAN: see

(a) `http://www.mathtools.net/MATLAB/Optimization/` for MATLAB and

(b) `http://www.nag.co.uk/numeric/fortran%5Flibraries.html` for NAg routines in FORTRAN and

(c) `http://www.sbsi-sol-optimize.com/manuals/SNOPT%20Manual.pdf`, and `http://www.sbsi-sol-optimize.com/asp/sol_product_snopt.htm`. for SNOPT (invented by Gill, Murray and Saunders) which may be called from a driver program in FORTRAN, or MATLAB;

Parameter	Significance of parameter	Low accuracy calculation	High accuracy calculation
In `Archi` `eps`	error-per-step tolerance in the ODE or DDE solver	10^{-6}	10^{-15}
In `LMDIF1` `epsfcn`	the relative errors in the objective functions $\approx$ `epsfcn`	default (0)	10^{-15}
`ftol`	relative change in the estimated minimum	10^{-6}	10^{-12}
`xtol`	relative change in the estimated parameter value	10^{-6}	10^{-12}

Table 2.3:

(d) `ftp://ftp.numerical.rl.ac.uk/pub/lancelot` for the FORTRAN package LANCELOT developed at Rutherford [103] (free to academic users, subject to conditions); see `http://www.numerical.rl.ac.uk/lancelot/blurb.html`.

Optimization procedures are generally iterative and the tolerances specified by the user govern the successful conclusion of an iterative process for determining the minimum of an objective function: (*i*) `ftol`[4],governing the relative change in the estimated minimum value of the objective function, (*ii*) `xtol`, governing the relative change in the argument at which the estimate of the minimum is attained. A parameter that is sometimes hidden (but can be of prime importance) is that specifying the accuracy to which the objective function is computed. In this respect, one may need to exercise caution about default parameter values. Thus, `LMDIF1` is a version[5] of `LMDIF` in which the calling sequence is simplified through the use of default parameter values (the user is required to specify fewer parameters). In consequence, a call to `LMDIF1` employs a default value (zero) of `epsfcn`. Our experience was transformed beyond recognition when the parameter `epsfcn` was set to an appropriate (non-default) value. Similar experience is likely with other codes. The calculations in our illustrations are based on the values in table 2.3.

Remark 2.4.1 Codes found in [109] are meant to facilitate parameter estimation. The code `Archi-L` includes a modified version of the optimization routine `LMDIF`; `Archi-N` invokes a NAg constrained minimization routine `E04UNF` [108] (in previous versions, `E04UPF`), to find the optimum parameter values.

[4] We use names employed in `LMDIF`.

[5] See `http://www.netlib.org/minpack/lmdif1.f`.

In our view[6], it is quite important to find (if possible) an optimization routine that, in searching for a solution to a constrained minimization problem, does not violate constraints that are essential to the code that evaluates the objective function. Thus, (*i*) DDE-solvers should not be expected to obtain a solution if τ is assigned a negative value, (*ii*) a problem in which the parameters can give rise to symptoms of stiffness may require different treatment from one that cannot. Where the optimizing routine does not have this feature (and more generally), it is helpful to have good initial estimates of the optimum parameters. These may be obtained from the optimum parameters for a simpler model in the hierarchy, or (a time-consuming process not given to automation) by plotting contour plots corresponding to parameters taken in pairs. Contour plotting *is* valuable when what is believed to be an optimum value has been located, as it allows a check to be made to ascertain the sensitivity of the objective function in the neighbourhood and to distinguish local and global minima.

Remark 2.4.2 A multitude of apparent local minima can be a consequence of too large an error tolerance in the ODE/DDE solver. (Compare the minimum of x^{2n} with the minima of $x^{2n} + \epsilon \sin(mx)$ where n, m are integers and ϵ corresponds to the maginitude of a perturbation.)

Contour plotting as an aid: potential problems

In practice, success in identifying parameters is related to the shape of the objective function, in terms of the parameters. The examples displayed here relate to actual models for the data presented.

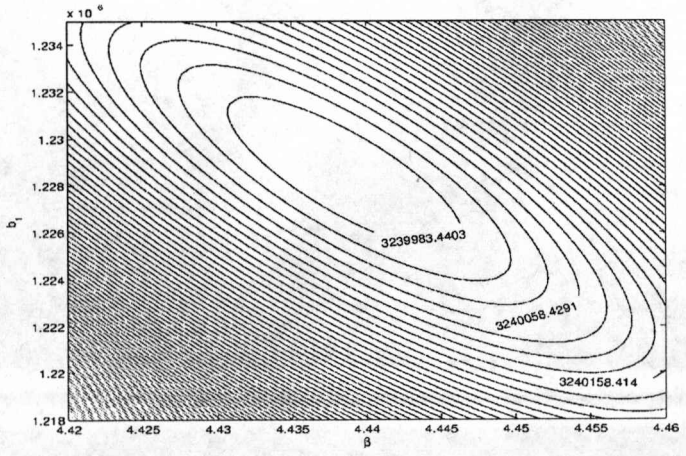


Figure 2.1: Contours suggesting that parameter estimation can be achieved successfully

[6] Others – we believe unwisely – take a more relaxed view.

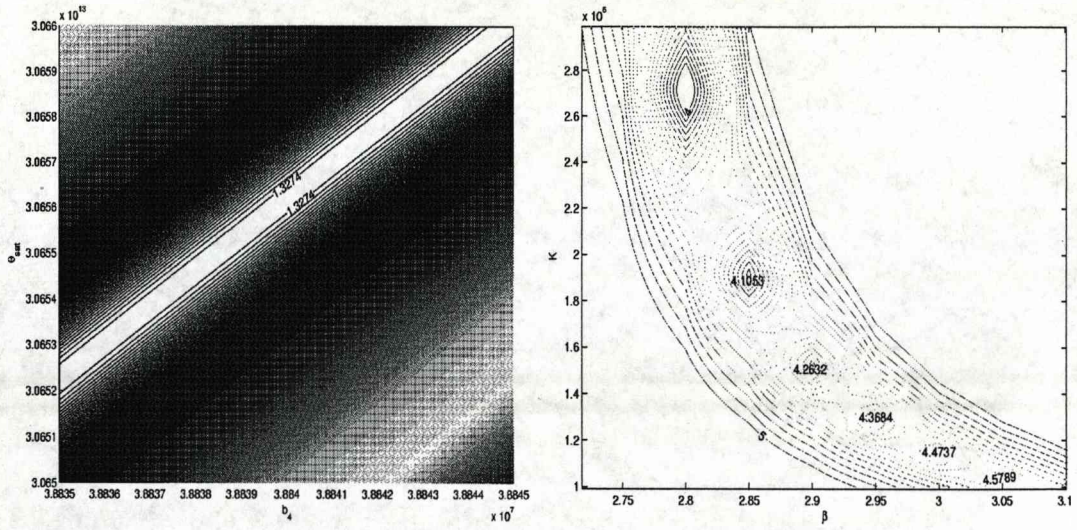


Figure 2.2: Contour plots.
Left: Model 4, log-least squares (vertical axis ϑ_{Sat}, horizontal axis b_4). Any combination of parameter components corresponding to the valley floor is equally effective at providing a small objective function.
Right: Model 1, ordinary least squares. Numerous local minima (which are spurious; see the text) cause a problem in the computation of best-fit parameter estimates. Vertical axis K, horizontal axis β.

A visual assessment can be obtained from graphical displays of the behaviour of $\Phi(p)$. Thus, successful minimization is relatively easily obtained when the objective function has a contour plot appearing like that in figure 2.1. However, a valley-type behaviour (figure 2.2, left-hand contour) implies a high correlation between corresponding parameters. The left-hand contour plot in figure 2.2 indicates that when ϑ_{Sat} and b_4 are appropriately related the objective function changes very little (the values of the residuals on all the included contours are identical to four decimal places). Actually, a large change in b_4 with a related large change in ϑ_{Sat} can result in a very small change in the objective function[7]. On the other hand, the right-hand contour in figure 2.2 shows a cluster of local minima near the global minimum. These apparent local minima are spurious, in the sense that they arise from using using too large a tolerance when solving the differential equations. Here, it is quite feasible that differing initial estimates of the parameters (employed to start the search for optimal values) could produce any one of the *local* minima.

[7]This example illustrates that the two parameters are not separately identifiable. The issue of a priori identifiability is addressed in [97].

2.4.4 Ordinary least squares objective function $\Phi_{OLS}(p)$

Parameter	M_1	M_2	M_3	M_4	M_5
β	4.44×10^{0}	4.36×10^{0}	4.52×10^{0}	4.52×10^{0}	4.50×10^{0}
K	3.99×10^{6}	3.23×10^{6}	3.17×10^{6}	3.17×10^{6}	3.19×10^{6}
γ	3.02×10^{-6}	3.48×10^{-6}	3.45×10^{-6}	3.48×10^{-6}	3.63×10^{-6}
b_i	1.23×10^{-6}	1.92×10^{0}	2.52×10^{0}	2.41×10^{0}	2.40×10^{0}
ϑ_{Sat}	–	2.46×10^{4}	1.34×10^{5}	1.31×10^{5}	1.15×10^{5}
τ	–	–	7.17×10^{-2}	8.98×10^{-2}	9.54×10^{-2}
α_E	0.0	9.14×10^{-2}	8.62×10^{-2}	9.1×10^{-2}	9.31×10^{-2}
T^*	–	–	–	1.24×10^{2}	1.40×10^{2}
r_m	–	–	–	–	5.17×10^{-3}
α_m	–	–	–	–	2.55×10^{-1}
Φ_{OLS}	1.05×10^{13}	4.49×10^{12}	4.04×10^{12}	3.91×10^{12}	3.78×10^{12}
$\breve{\mu}_{cAIC}$	472.2	467.0	475.4	488.9	544.4

Table 2.4: Best-fit parameter estimates for ordinary least-squares and the corrected Akaike indicator. Calculations are based on `LMDIF` and the values `eps` $= 10^{-6}$, `ftol` $= 10^{-6}$, `xtol` $= 10^{-6}$, `epsfcn` $= 0$.

Parameter estimation results obtained using ordinary least squares approach for Models 1 to 5 are summarized in table 2.4. An increase in the number of model parameters provides a better description of the data in terms of the minimized value of the objective function. However, the increasing values of the corrected Akaike index indicate a gradual information loss for the given data set, as the complexity of models increases. Variation of the best-fit parameter estimates between the models is within $\pm 10\%$, except for the estimate of ϑ_{Sat}. Further, the data set does not provide a biologically correct estimate of the time lag of cell division τ. (Rather, the delay estimate obtained via ordinary least squares corresponds to a realistic duration of some stage of the cell cycle.) Visual inspection of graphs of $V(t)$ and $E(t)$ suggests that Model 1 nicely approximates the viral load data, but rather poorly approximates the CTL data. The other models describe much better the CTL kinetics at the expense of a somewhat poorer agreement with the virus data ($V(t)$).

2.4.5 Weighted least squares objective function $\Phi_{\Omega LS}(p)$

We used $\Phi_{\Omega LS}(p) = \sum_{j=1}^{N} \sum_{i=1}^{M} \left\{ \omega_i^{[j]} \left[y^i(t_j, p) - y_j^i \right] \right\}^2$ with the weights $\omega_i^j = \{\overline{y}_j^i\}^{-1}$ obtained using the means $\overline{y}_j^i$ of the data in table 2.1. The best-fit parameter estimates are summarized in table 2.5. Again, an increase in the number of model parameters reduces bias in the description of the data but is associated with an increasing value of the Akaike criterion. Interestingly, the weighted least

squares approach ensured biologically correct estimates of the mean cell division time, the delay τ ranging from 5 to 7 hours. Qualitatively, Models 1 to 4 provide a similar approximation of the viral load data whereas model 5 fits the CTL data somewhat better than the other models from the set

Parameter	Model 1, M_1	Model 2, M_2	Model 3, M_3	Model 4, M_4	Model 5, M_5
β	4.68×10^0	4.67×10^0	4.60×10^0	4.59×10^0	4.60×10^0
K	2.80×10^6	2.79×10^6	2.86×10^6	2.83×10^6	3.04×10^6
γ	5.19×10^{-6}	7.29×10^{-6}	9.30×10^{-6}	9.99×10^{-6}	1.48×10^{-5}
b_i	1.09×10^{-6}	2.31×10^0	2.48×10^0	2.02×10^0	2.03×10^0
ϑ_{Sat}	–	3.12×10^5	1.90×10^5	9.65×10^4	4.64×10^4
τ	–	–	0.291×10^0	0.207×10^0	0.276×10^0
α_E	0.0	10^{-32}	0.0	3.30×10^{-7}	2.2×10^{-2}
T^*	–	–	–	8.27×10^2	8.22×10^2
r_m	–	–	–	–	8.73×10^{-4}
α_m	–	–	–	–	1.60×10^{-4}
$\Phi_{\Omega LS}$	6.36×10^0	6.06×10^0	5.99×10^0	4.95×10^0	4.43×10^0
$\breve{\mu}_{cAIC}$	50.2	57.0	66.9	78.0	132.3

Table 2.5: Best-fit parameter estimates for weighted least-squares and the corrected Akaike indicator. Calculations are based on `LMDIF` and `eps` $= 10^{-6}$, `ftol` $= 10^{-6}$, `xtol` $= 10^{-6}$, `epsfcn` $= 0$.

The best-fit values of the objective function suggest (in each case) that the maximum likelihood estimate of the data variance is about $\widehat{\sigma}^2 \approx 0.3$ to 0.4. $(\widehat{\sigma}^2 = \frac{1}{NM}\sum_j \|diag^{-1}[\omega_1^{[j]}, \omega_2^{[j]}, \ldots, \omega_M^{[j]}][y(t_j, \widehat{p}) - y_j]\|^2 = \frac{1}{NM}\Phi_{\Omega LS}(\widehat{p}))$.

2.4.6 Log-least squares objective function $\Phi_{LogLS}(p)$

Parameter estimation results obtained using log-least squares approach for Models 1 to 5 are shown in table 2.6. (We fit to data summarized as the means of the log-values rather than the log of the means.) An increase in the number of parameters does not affect the bias in the data description by Models 1 to 3 but it does provide a better description of the data when applying Models 4 and 5. The increasing value of the corrected Akaike indicator suggests a gradual information loss for the given data set as the complexity of the models increases. The log-least squares strategy fails to identify reliably the parameters b_i and ϑ_{Sat}, as these appear to be highly correlated (see fig. 2.2). From another viewpoint, the estimates of r_m, α_m, the memory CTL turnover, appear to be biologically more consistent than those obtained with ordinary- or weighted least squares. Again, the strategy does not support the need to consider a delay in the model for the given data set.

Parameter	M_1	M_2	M_3	M_4	M_5
β	5.49×10^{0}	5.49×10^{0}	5.49×10^{0}	4.70×10^{0}	4.70×10^{0}
K	9.45×10^{5}	9.45×10^{5}	9.45×10^{5}	1.31×10^{6}	1.31×10^{6}
γ	2.12×10^{-6}	2.12×10^{-6}	2.12×10^{-6}	2.25×10^{-6}	2.24×10^{-6}
b_i	2.05×10^{-6}	4.78×10^{6}	4.78×10^{6}	3.88×10^{7}	3.88×10^{7}
ϑ_{Sat}	–	2.33×10^{12}	2.33×10^{12}	3.07×10^{13}	3.07×10^{13}
τ	–	–	0.00	0.00	5.60×10^{-3}
α_E	2.35×10^{-2}	2.35×10^{-2}	2.35×10^{-2}	1.39×10^{-2}	1.31×10^{-2}
T^*	–	–	–	1.345×10^{3}	1.343×10^{3}
r_m	–	–	–	–	1.10×10^{-2}
α_m	–	–	–	–	9.10×10^{-3}
Φ_{LogLS}	2.443×10^{0}	2.443×10^{0}	2.443×10^{0}	1.761×10^{0}	1.764×10^{0}
$\breve{\mu}_{cAIC}$	35.9	43.4	53.4	62.5	118.5

Table 2.6: Best-fit parameter estimates for log least-squares and the corrected Akaike indicator. Calculations were based on `LMDIF` and eps $= 10^{-6}$, `ftol` $= 10^{-6}$, `xtol` $= 10^{-6}$, `epsfcn` $= 0$.

All the models with parameters estimated following the log-least squares strategy seem to capture the overall kinetics of the data quite well.

2.4.7 Computing with greater accuracy

Parameter	M_1	M_2	M_3	M_4	M_5
β	4.61×10^{0}	4.51×10^{0}	4.62×10^{0}	4.61×10^{0}	4.61×10^{0}
K	2.70×10^{6}	4.69×10^{6}	5.01×10^{6}	4.98×10^{6}	5.07×10^{6}
γ	1.39×10^{-6}	8.04×10^{-5}	3.29×10^{-4}	2.96×10^{-4}	2.45×10^{-4}
b_i	9.22×10^{-7}	1.42×10^{0}	1.14×10^{0}	1.16×10^{0}	1.22×10^{0}
ϑ_{Sat}	–	0	8.79×10^{-6}	4.59×10^{-6}	2.45×10^{-4}
τ	–	–	4.38×10^{-2}	4.15×10^{-2}	4.38×10^{-2}
α_E	9.29×10^{-2}	2.01×10^{-1}	1.02×10^{-1}	1.02×10^{-1}	1.03×10^{-14}
T^*	–	–	–	1.09×10^{0}	134×10^{0}
r_m	–	–	–	–	2.12×10^{-1}
α_m	–	–	–	–	2.20×10^{-1}
Φ_{OLS}	6.54×10^{12}	7.82×10^{11}	1.60×10^{12}	1.60×10^{12}	1.37×10^{12}
$\breve{\mu}_{cAIC}$	465.1	440.8	461.5	475.5	529.2

Table 2.7: Best-fit parameter estimates for ordinary least-squares and the corrected Akaike indicator when the higher accuracy numerical solution is used: Calculations were based on `Archi-N` with `E04UNF`, and eps $= 10^{-15}$, `ftol` $= 10^{-12}$, `xtol` $= 10^{-12}$, `epsfcn` $= 10^{-15}$.

Parameter	M_1	M_2	M_3	M_4	M_5
β	5.14×10^{0}	4.56×10^{0}	4.57×10^{0}	4.60×10^{0}	4.60×10^{0}
K	1.23×10^{5}	4.42×10^{6}	4.46×10^{6}	3.27×10^{6}	3.26×10^{6}
γ	1.90×10^{-6}	5.47×10^{-5}	7.93×10^{-5}	2.14×10^{-5}	2.23×10^{-5}
b_i	1.37×10^{-5}	1.41×10^{0}	1.41×10^{0}	2.34×10^{0}	2.41×10^{0}
ϑ_{Sat}	–	0	3.49×10^{-13}	4.854×10^{4}	5.39×10^{4}
τ	–	–	1.76×10^{-2}	5.18×10^{-1}	5.04×10^{-1}
α_E	2.09×10^{-2}	1.09×10^{-1}	1.09×10^{-1}	1.09×10^{-1}	2.52×10^{-10}
T^*	–	–	–	1.663×10^{3}	1.508×10^{3}
r_m	–	–	–	–	1.32×10^{-1}
α_m	–	–	–	–	5.10×10^{-1}
$\Phi_{\Omega LS}$	5.23×10^{0}	5.36×10^{0}	5.15×10^{0}	4.18×10^{0}	4.18×10^{0}
$\breve{\mu}_{cAIC}$	47.3	55.2	64.6	75.5	131.5

Table 2.8: Best-fit parameter estimates for weighted least-squares and the corrected Akaike indicator when the higher accuracy numerical solution is used: Calculations were based on `Archi-N` with `E04UNF`, and `eps` $= 10^{-15}$, `ftol` $= 10^{-12}$, `xtol` $= 10^{-12}$, `epsfcn` $= 10^{-15}$.

The calculations represented in table 2.4 were checked by refining the parameters that govern the accuracy. The original values `eps` $= 10^{-6}$, `ftol` $= 10^{-6}$, `xtol` $= 10^{-6}$, `epsfcn` $= 0$ were replaced by `eps` $= 10^{-15}$, `ftol` $= 10^{-12}$, `xtol` $= 10^{-12}$, `epsfcn` $= 10^{-15}$, and the figures in table 2.7 were then obtained. The refined tolerances have a noticeable effect on the parameter values and in consequence on the ranking of the parametrized models. Models with parameters computed with lower tolerance are ranked with respect to the information loss in the ranking :

$$M2 \text{ (best) - } M1 \text{ - } M3 \text{ - } M4 \text{ - } M5;$$

with parameters computed to the higher tolerance we obtain the ranking

$$M2 \text{ (best) - } M3 \text{ - } M1 \text{ - } M4 \text{ - } M5.$$

In both cases Model 2 has the least Akaike index. The parameter ϑ_{Sat} (which represents the viral load for half-maximal CTL-stimulation) does not occur in Model 1; in the high-accuracy figures for Model 2 it is close to zero. Such a small value represents an effectively immediate response to the infection (what is considered to be a "programmed" response, in [104]), irrespective of the viral load.

If we use Model 5, the data does not support a biologically correct estimation of the memory cell life-span α_m. We note that in a similar parameter estimation study [104] the parameter α_m was assigned a value rather than estimated.

Table 2.9 summarizes the analysis of the confidence in the best-fit parameter estimates for Models 1 to 3 using high accuracy solutions and following OLS.

The ranking of the models according to the Akaike criteria suggests that the least information loss should be the feature of Model 2 as compared to Models 1 or 3. The reader may compare the widths of the confidence intervals for the same parameter in differing models, displayed in table 2.9. Other things being equal, a narrower interval is to be preferred. An unduly large confidence interval indicates that the parameter is unidentifiable for practical purposes.

Parameter	M_1	M_2	M_3
β	4.61×10^{0}	4.51×10^{0}	4.62×10^{0}
95% CI :	$[4.00 \times 10^{0}, 5.23 \times 10^{0}]$	$[4.23 \times 10^{0}, 4.76 \times 10^{0}]$	$[4.43 \times 10^{0}, 4.85 \times 10^{0}]$
K	2.70×10^{6}	4.69×10^{6}	5.01×10^{6}
95% CI :	$[2.28 \times 10^{6}, 3.00 \times 10^{6}]$	$[4.20 \times 10^{6}, 5.20 \times 10^{6}]$	$[4.62 \times 10^{6}, 5.45 \times 10^{6}]$
γ	1.39×10^{-6}	8.04×10^{-5}	3.29×10^{-4}
95% CI :	$[1.17 \times 10^{-6}, 1.71 \times 10^{-6}]$	$[7.54 \times 10^{-5}, 8.58 \times 10^{-5}]$	$[3.20 \times 10^{-4}, 3.31 \times 10^{-4}]$
b_i	9.22×10^{-7}	1.42×10^{0}	1.141×10^{0}
95% CI :	$[7.96 \times 10^{-7}, 1.01 \times 10^{-6}]$	$[1.40 \times 10^{0}, 1.43 \times 10^{0}]$	$[1.134 \times 10^{0}, 1.143 \times 10^{0}]$
ϑ_{Sat}	–	0 (3.23×10^{-176})	8.79×10^{-6}
95% CI :	–	$[0, 1.2 \times 10^{-164}]$ If $\vartheta_{Sat} = 0$, (2.16) results.	$[5.25 \times 10^{-6}, \vartheta^{\max}]$ where $\vartheta^{\max} \geq 8 \times 10^{-5}$.
τ	–	–	4.38×10^{-2}
95% CI :	–	–	$[4.24 \times 10^{-2}, 4.43 \times 10^{-2}]$
α_E	9.29×10^{-2}	2.01×10^{-1}	1.02×10^{-1}
95% CI :	$[4.84 \times 10^{-2}, 1.73 \times 10^{-1}]$	$[1.19 \times 10^{-1}, 2.14 \times 10^{-1}]$	$[1.01 \times 10^{-1}, 1.15 \times 10^{-1}]$
Φ_{OLS}	6.54×10^{12}	7.82×10^{11}	1.6×10^{12}
$\breve{\mu}_{cAIC}$	465.1	440.8	461.5

Table 2.9: Estimates of 95% confidence intervals, for the best-fit high accuracy parameter estimates for Models 1 to 3 using ordinary least-squares. Estimates were computed using `Archi-N` (see Remarks 2.4.1 & 2.4.3), with `E04UNF`, eps $= 10^{-15}$, `ftol` $= 10^{-12}$, `xtol` $= 10^{-12}$, `epsfcn` $= 10^{-15}$ as in table 2.7.

Remark 2.4.3 The evaluation of confidence intervals for parameter estimates is easily described but can be computationally difficult.

- The process is often computationally highly expensive. This is particularly the case where the interval is large; asking for 67% confidence rather than 95% confidence will reduce the length of the interval.
- We have found little difficulty when using the NAg minimization routine `E04UNF` [108] (apart from the demands for computer time), provided the positivity constraints are incorporated as indicated in Remark 2.3.1. The use of an output parameter `IFAIL` in `E04UNF` allows monitoring of "soft failure" as well as "hard failure". The numbers in table 2.9 were obtained with this optimization code, in `Archi-N`. We recalculated those estimates

using `LMDIF1` in `Archi-L`, and obtained the same optimum parameters but somewhat different confidence intervals.

- The optimization codes proceed iteratively and the starting estimate may affect the outcome. Confidence intervals are produced by investigating a range of values p_ℓ and optimizing over the remaining parameters. It is a reasonable strategy to take the optimum parameters for p_ℓ as a starting estimate for the optimum values for $p_\ell + \delta p_\ell$. Values in table 2.9 were obtained with this strategy. However one can envisage situations where the optimization code was not sufficiently robust to compute a *global* minimum. Further investigation into robust optimization codes, and their useage, appears to be suggested.

The Akaike indices suggest that the best model is Model 2 (M_2). One might therefore anticipate that M_2 would have the smallest confidence intervals for the best-fit parameter estimates. In reality, it appears that M_3 is characterized by smaller confidence intervals. This might be a consequence of a number of factors, for example (*i*) the Akaike indices provide approximate indicators, (*ii*) one needs to check whether the differences in the estimated Akaike indicators are statistically significant, given the errors in the data.

As stated above, the values of the Akaike indicators suggest that for the given data set the model which ensures the least information loss is Model 2:

$$\frac{d}{dt}V(t) = \beta \cdot V(t) \cdot \left(1 - \frac{V(t)}{K}\right) - \gamma \cdot V(t) \cdot E(t) \tag{2.14}$$

$$\frac{d}{dt}E(t) = b_2 \cdot \frac{V(t)}{\vartheta_{Sat} + V(t)} \cdot E(t) - \alpha_E \cdot E(t) \tag{2.15}$$

For this model, the best-fit estimate for the viral load necessary for the half-maximal CTL division rate (ϑ_{Sat}) is close to zero. If $\vartheta_{Sat} = 0$ and $V(t) \neq 0$, (2.15) becomes

$$\frac{d}{dt}E(t) = b_2 \cdot E(t) - \alpha_E \cdot E(t). \tag{2.16}$$

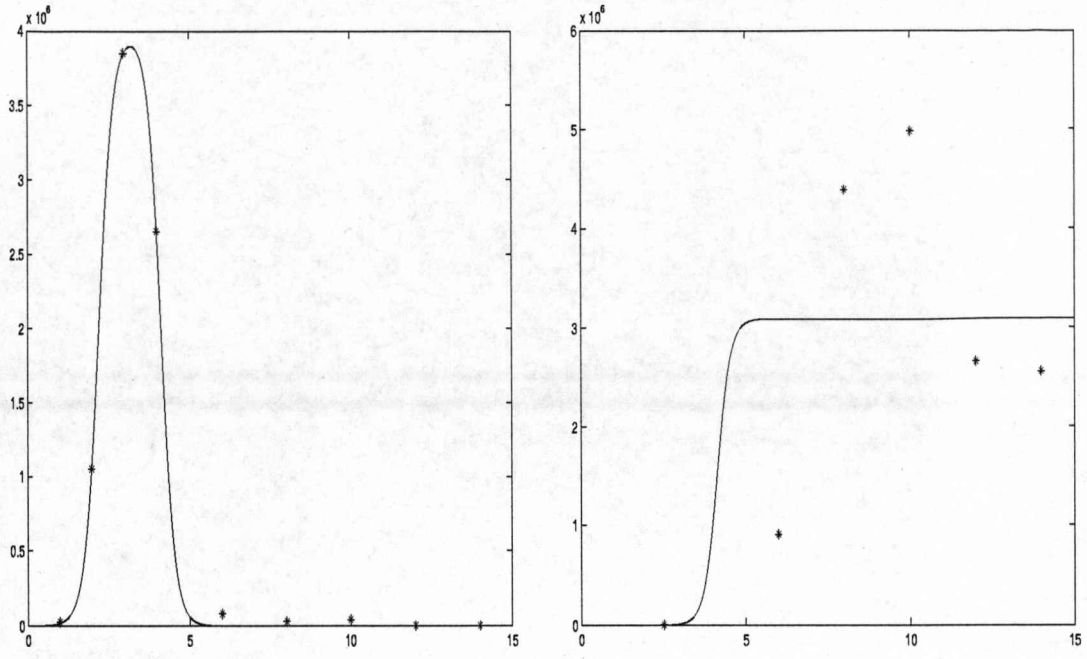
As long as a virus population is present in the host, the second equation of Model 2 can be replaced by the simplified linear form:

$$\frac{d}{dt}E(t) = \underbrace{b_2 \cdot H_{V^{\vartheta_{Sat}}} \cdot E(t)}_{\text{Modification suggested by the best-fit estimate}} - \alpha_E \cdot E(t), \tag{2.17}$$

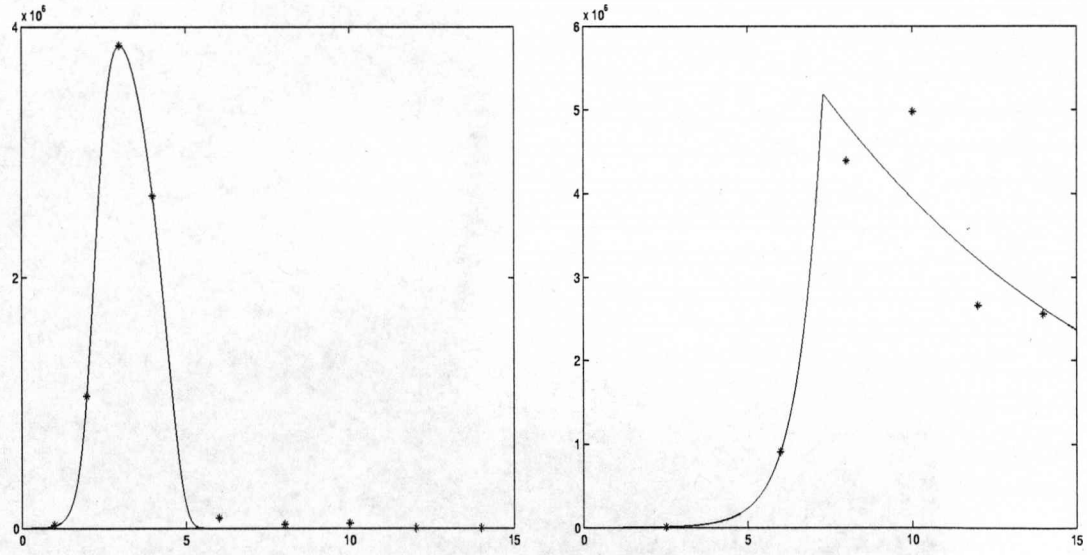
where the cell proliferation term on the right-hand side uses the Heaviside function

$$H_{V^{\vartheta_{Sat}}} = 0 \text{ for } V \leq V^{\vartheta_{Sat}};\ H_{V^{\vartheta_{Sat}}} = 1 \text{ for } V \geq V^{\vartheta_{Sat}};$$

as a characteristic value for $V^{\vartheta_{Sat}}$ one can take $V^{\vartheta_{Sat}} \approx 1$. It is remarkable that the best-approximating model for the typical data set for the LCMV-CTL population dynamics in primary infection appears to be the one which was introduced elsewhere in an *ad hoc* manner (see [104]). Biologically, the form of the proliferation term implies that the CTL response to a low-dose LCMV infection is a process regulated by the virus load in an "on" (full activation) and "off" (no activation at all) way."



$V(t)$ (*a*) Model 1 (see caption) $E(t)$



$V(t)$ (*b*) Model 2 (see caption) $E(t)$

Figure 2.3: Ordinary least squares.
(*a*) Model 1 computed using low accuracy (table 2.4);
(*b*) Model 2 computed using high accuracy (table 2.7).
Shown are the best-fit predictions (solid lines) for the viral load, $V(t)$ and for the number of CTLs, $E(t)$, and the mean values (* symbols) for the original data. *Horizontal axes:* $0 \leq t \leq 15$; *vertical axes:* $0 \leq V(t) \leq 4 \times 10^6$; $0 \leq E(t) \leq 6 \times 10^6$.

Chapter 3

Introduction to stochastic processes

3.1 Brownian motion and MATLAB simulations

In this section we will introduce the idea of Brownian motion and investigate the simulation of discretized Brownian paths using MATLAB.

3.1.1 Brownian motion

In 1828 the Scottish botanist Robert Brown described the motion of a pollen particle suspended in fluid. He observed that a particle moved in a random irregular fashion and from this he obtained the equations for what has been called Brownian motion [58, 73]. Louis Bachelier (1900), Albert Einstein (1905) and Norbert Wiener (1923) began developing the theory of Brownian motion [66]. Some texts refer to Brownian motion as the physical process and refer to the mathematical process as a Wiener process [58].

A clear, precise definition of a scalar Brownian motion is given in Higham [49], and is quoted below.

"A scalar standard Brownian motion, or standard Wiener process, over $[0, T]$ is a random variable $W(t)$ that depends continuously on $t \in [0, T]$ and satisfies the following three conditions:

1. $W(0) = 0$ (with probability 1).
2. For $0 \leq s < t \leq T$ the random variable given by the increment $W(t) - W(s)$ is normally distributed with mean zero and variance $t - s$; equivalently, $W(t) - W(s) \sim \sqrt{t-s}N(0,1)$, where $N(0,1)$ denotes a normally distributed random variable with zero mean and unit variance.
3. For $0 \leq s < t < u < v \leq T$ the increments $W(t) - W(s)$ and $W(v) - W(u)$ are independent."

To extend our investigations to SDDEs we will need to simulate discretized Brownian paths of $W(t)$, specified at discrete t values. We will divide our interval $[0, T]$ into N equal intervals of width $h = T/N$. With the usual notation we define $W_i = W(t_i)$ where $t_i = ih$. Condition 1 of the definition gives $W_0 = 0$. Condition 2 gives us the recurrence $W_i = W_{i-1} + dW_i$, $i = 1, 2, \ldots, N$, where each dW_i is a random variable of the form $\sqrt{h}N(0, 1)$, and condition 3 states that the dW_i are independent. To simulate our path we need to generate N independent values from the $N(0, 1)$ distribution. For this we can use MATLAB's random number generator command "*randn*('*state*', n)". This command has the advantage of making experiments repeatable. If we set $n = 1$, *randn* will produce the same sequence of random variables at each run of the program. Setting n to some other integer value will produce a different repeatable sequence. We can now scale our random values by $\sqrt{h}$ to simulate the N values of dW_i. The remainder of this thesis depends upon the production of 500 simulated Brownian paths so it would be prudent at this point to discuss the process by which MATLAB generates *randn* and to check that the command actually produces values from the $N(0, 1)$ distribution.

3.1.2 MATLAB's *randn*('*state*', n) command

Marsaglia's ziggurat algorithm with a period of approximately 2^{64} is used to generate *randn*('*state*', s), with s taking any integer value from 0 to $2^{32} - 2$ to initiate the algorithm. See MATLAB's "help" toolbar for this information.

The Marsaglia Ziggurat Algorithm

We give here a simplified description of the algorithm.

The area of the right hand half of $y = e^{-x^2/2}$, the non-standardised shape of $N(0, 1)$, is split into a large number of horizontal strips of equal area. The figure 3.1 shows this split for just 3 strips, but in practice 128 or 256 are used. A random y coordinate can be selected using the uniform distibution by $y = U(0, 1)$. This y coordinate selects the horizontal strip, S_i, say. If we define the lower boundary of this strip by the interval $[0, x_i]$ then another selection from the uniform $U(0, 1)$ distribution supplies our random $N(0, 1)$ variate as $x = x_i.U$. However, it is possible that the point (x, y) is not contained within the density curve, especially if x is close to x_i. This occurs when the point (x, y) is in the top righthand corner of a strip, outside of the curve. The labels B_1, B_2 in figure 3.1 are actually placed in such positions. In this case our value is rejected and the method is repeated. The method is defined in more detail in Marsaglia & Tsang [65].

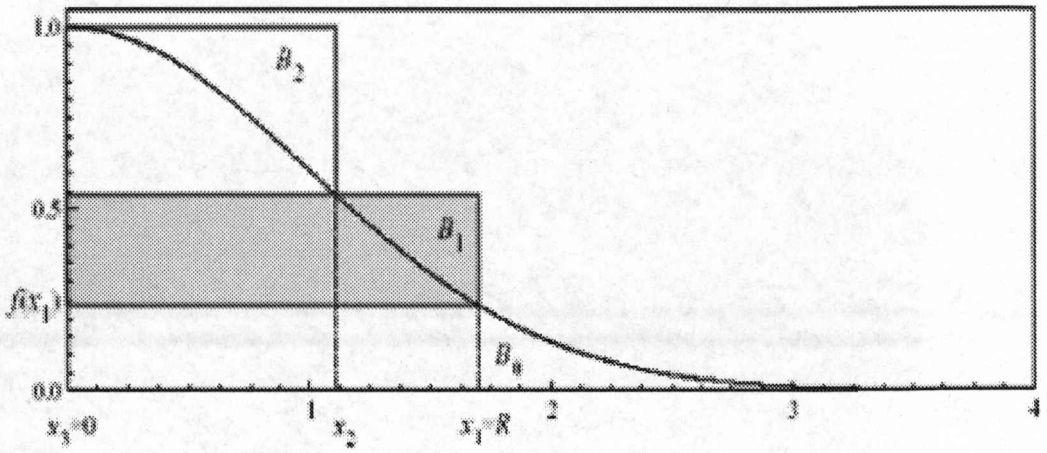


Figure 3.1: Example of a three-way ziggurat partitioning of the normal density curve, taken from [27].

Validation of *randn* output

Recent papers [27, 62] have questioned the efficiency of the algorithm. In [27, 62] it is suggested that an improved version, using a different random number generator, passes statistical tests for normality. With some doubt expressed about the ziggurat algorithm it would seem prudent to conduct experimental work on *randn* to confirm that we do actually get normal variates for our later experiments. We used a simple MATLAB code, listed in appendix B, to produce vectors containing 10000 and 100000 values from *randn*. These values were chosen as they represent the minimum and maximum vector sizes for our following experiments on SDDEs. Figures 3.2, 3.3 show histograms of the values, together with superimposed standard normal curves. Both show that the histograms are very good approximations to the expected standard normal curve.

As a further check a similar figure 3.4 was produced using a vector of 1000000 values and we can see that the histogram is a nearly perfect fit to the standard normal curve.

We used same code to check the mean and standard deviation of the vectors of values from *randn*. For the more relevant vectors of size 10000, 100000 we produced five vectors using different settings in the *state* statement. However, to consider the limiting case as the magnitude of the vector increases we also produced the parameter values for vectors of length 1000000 and 10000000. All of the experimental results are given in table 3.1.

The table clearly confirms the results we are hoping for. The mean values are all very close to zero and the standard deviation values are very close to

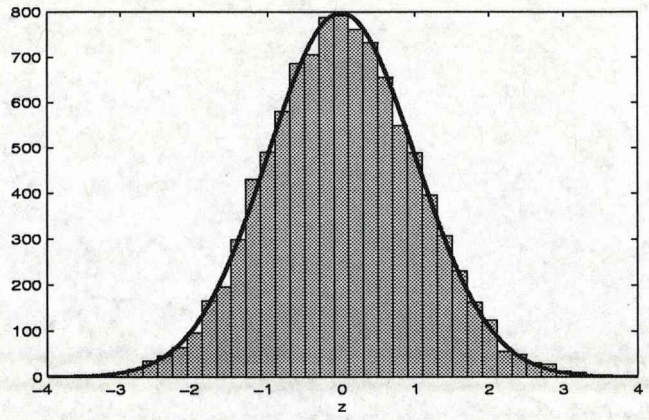


Figure 3.2: Histogram of 10000 values from *randn*, with superimposed normal distribution curve

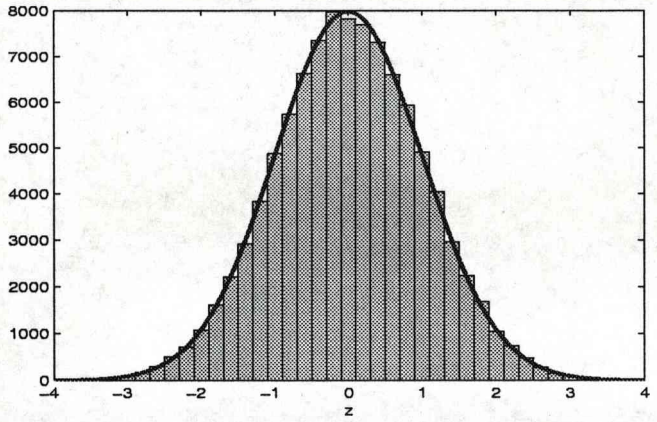


Figure 3.3: Histogram of 100000 values from *randn*, with superimposed normal distribution curve

unity, with these approximations improving as the vector length increases. We can conclude that *randn* will give us good variates from the standard normal distribution.

3.1.3 Brownian paths using *randn*

We can now confidently use *randn* with MATLAB to simulate the Brownian paths required for the remaining chapters of this thesis. Using the coding listed in appendix B, we have produced figure 3.5 showing a single path and figure 3.6 showing ten paths. Finally, figure 3.7 shows 20 paths together with their mean

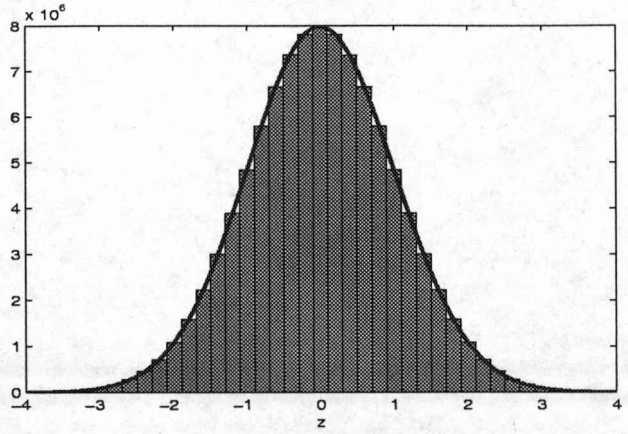


Figure 3.4: Histogram of 1000000 values from *randn*, with superimposed normal distribution curve

Vector length	"state"	Mean	Standard deviation
10000	1	0.007456	1.0111
10000	2	-0.000655	0.9962
10000	3	-0.012478	1.0021
10000	4	0.001682	0.9962
10000	5	-0.002675	0.9901
100000	1	0.005173	1.0014
100000	2	0.001578	1.0002
100000	3	-0.004220	1.0012
100000	4	0.001380	0.9983
100000	5	-0.003255	0.9965
1000000	1	0.000857	0.99979
10000000	1	0.000050	1.00023

Table 3.1: Mean and Standard deviation values

value, in red, at each point in time. As we would expect, the mean value approximates to the line $y = 0$.

3.2 Stochastic calculus

Sample Brownian paths are nowhere differentiable and have unbounded variation, and this has major consequences for the definition of the stochastic integral with respect to Brownian paths, [66]. In this section we will give a brief summary of the process of finding a stochastic integral. We refer the reader

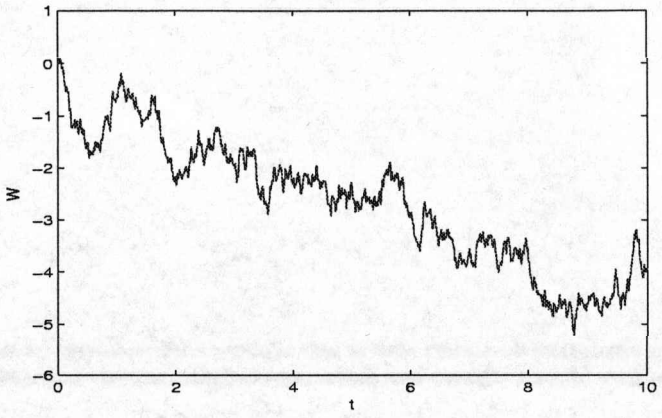


Figure 3.5: Single Brownian path

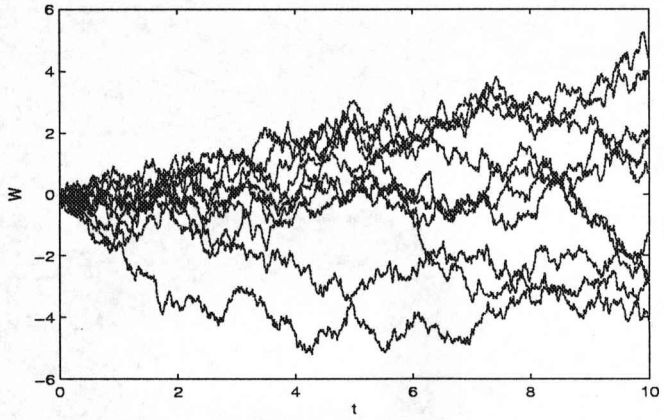


Figure 3.6: Ten Brownian paths

to [57, 58, 66, 73, 78, 75] for a more rigorous approach. If we consider the integral $\int_a^b f(t)dt$ then we can partition $[a, b]$ by $a = t_0 < t_1 < \ldots < t_n = b$. If we let $\delta t_i = t_i - t_{i-1}$ then we can define the Riemann sum

$$S_n = \sum_{i=1}^{n} f(x_i)\delta t_i$$

where $x_i \in [t_{i-1}, t_i]$. If the limit exists,

$$\lim_{n\to\infty} S_n = \int_a^b f(t)dt.$$

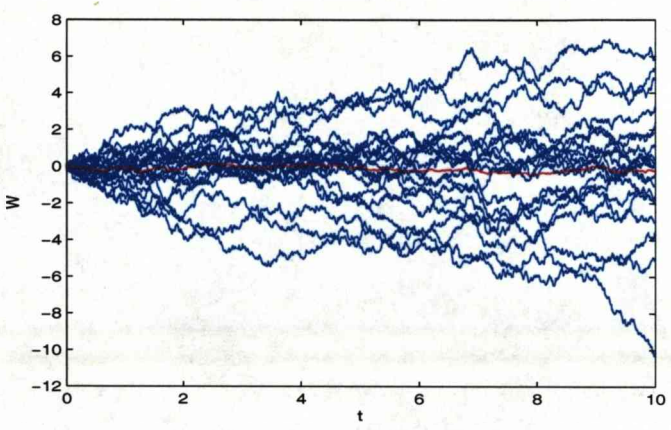


Figure 3.7: Twenty Brownian paths, with mean path (in red)

We would now like to be able to define $\int_0^1 f(t)dW(t)$ where $f(t)$ is a function or a stochastic process on $[0,1]$ and $W(t)$ is a Brownian path. Unfortunately, $W(t)$ does not have a derivative and so we cannot write our integral as a Riemann integral, [73]. However, the Riemann-Stieltjes integral allows us to find a kind of pathwise integral, [66]. Mikosch [66] states that we can evaluate integrals of the form $\int_0^1 e^t dW(t)$, but not explicitly in terms of Brownian motion. A major problem is that Brownian paths do not have bounded variation and hence the Riemann-Stieltjes integral approach must fail, [58].

3.2.1 Definitions of *Itô* and Stratonovich integrals

The following definitions are taken from [73]. If we define the indicator function by

$$\chi_{[t_i,t_{i+1}]}(t) = \begin{cases} 1 & \text{when } t \in [t_i, t_{i+1}], \\ 0 & \text{otherwise.} \end{cases}$$

and take $t_i^* \in [t_i, t_{i+1}]$ then we can approximate f(t) by

$$\sum_i f(t_i^*).\chi_{[t_i,t_{i+1}]}(t).$$

We can now define $\int_0^1 f(t)dW(t)$ as the limit

$$\sum_i f(t_i^*)[W_{t_{i+1}} - W_{t_i}] \qquad as\ n \to \infty.$$

We note that if we take

1. $t_i^* = t_i$ then we have defined the *Itô* integral, while taking

2. $t_i^* = \frac{t_i + t_{i+1}}{2}$ gives us the Stratonovich integral.

The most common example used in texts is $\int_0^T W dW$ (for example, [58, 66, 73]). It can be shown that in the *Itô* definition this integral is $\frac{1}{2}(W^2(T) - T)$, whereas with the Stratonovich definition gives $\frac{1}{2}W^2(T)$, as one would expect from ordinary calculus. The difference between these two integrals follows from the lack of smoothness of $W(t)$, and can be explained by *Itô*'s stochastic chain rule formula, [73]. The choices of t_i^* for the two integrals defined above are the most useful. The *Itô* integral is a martingale and the Stratonovich gives the results expected from ordinary calculus, [58]. It is to be noted that the stochastic integrals depend upon the particular Brownian path. We will investigate stochastic integrals using Matlab in the next section.

3.3 Stochastic integral simulations

Using *randn* we simulated a Brownian path with increment 0.00005. First we used this path to estimate both the *Itô* and the Stratonovich integrals of $\int_0^1 W dW$ with δt decreasing from 1 to 0.00001 and hence n increasing from 1 to 10000. We also used two different simulated Brownian paths. Tables 3.2, 3.3 show the results for the two integration definitions. Finally, table 3.4 shows the *Itô* and Stratonovich mean values for $\int_0^1 W dW$ over 10000 paths for various step sizes. As expected the difference between the integrals is approaching the exact difference of 0.5 as the step size decreases.

Steps	$randn('state', 1)$	$randn('state', 200)$
1	0	0
10	-0.30770	-0.54033
100	-0.46139	-0.43228
1000	-0.46442	-0.33178
10000	-0.46368	-0.32890
$W(1)$	-0.28442	-0.61642
$\frac{1}{2}(W^2(1) - 1)$	-0.45955	-0.31001

Table 3.2: *Itô* values for $\int_0^1 W dW$ with exact values, $\frac{1}{2}(W^2(1) - 1)$

It is interesting to note that Higham, in [49], actually regards the Riemann sum as an approximation to the stochastic integral. We will also see that the numerical methods used later will adopt this approach. We can investigate this by using the definitions to estimate both the *Itô* and the Stratonovich integrals of $\int_0^1 t dW$ for one of the sample paths used in the previous experiment. Table 3.5 shows the results for our particular path. As $f(t) = t$ is a smooth increasing

Steps	$randn('state', 1)$	$randn('state', 200)$
1	-0.13246	0.24761
10	0.22114	0.03972
100	0.12990	0.27099
1000	0.02547	0.16020
10000	0.04440	0.18659
$W(1)$	-0.28442	-0.61642
$\frac{1}{2}W^2(1)$	0.04048	0.18999

Table 3.3: Stratonovich values for $\int_0^1 W dW$ with exact values, $\frac{1}{2}W^2(1)$

Steps	*Itô* mean	Stratonovich mean	*Itô* standard deviation	Stratonovich standard deviation
1	0	0.49543	0	0.8487
10	-0.00624	0.49210	0.6424	0.6966
100	-0.00954	0.49010	0.6750	0.6803
1000	-0.01069	0.48916	0.6782	0.6787
10000	-0.01085	0.48918	0.6786	0.6787

Table 3.4: $\int_0^1 W dW$ using 10000 paths

Steps	*Itô*	Stratonovich
1	0	-0.14299
10	-0.07325	-0.08755
100	-0.07612	-0.07755
1000	-0.07395	-0.07409
10000	-0.07381	-0.07382

Table 3.5: $\int_0^1 t dW$ using path $randn('state', 1)$

function we find that the Stratonovich value is marginally greater than the *Itô* value. The tables indicate that the integrals are converging. We also calculated the integral with 10000 different paths and table 3.6 shows the mean and standard deviation for both types of integral using the same step sizes.

3.3.1 Conclusion

The experiments conducted using MATLAB and its random normal variate generator indicate that we can be fairly confident in using these features in our future experimental work.

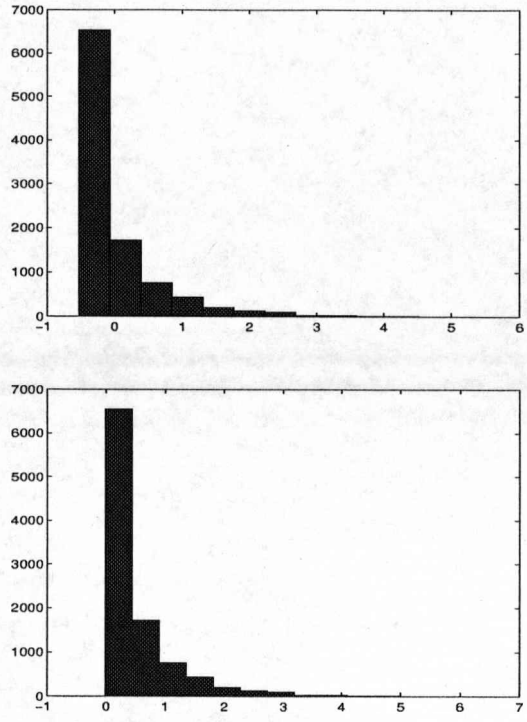


Figure 3.8: Histogram of values of $\int_0^1 W dW$ using 10000 paths
Top: *Itô* integral Bottom: Stratonovich integral

Steps	*Itô* mean	Stratonovich mean	*Itô* standard deviation	Stratonovich standard deviation
1	0	-0.006193	0	0.4946
10	-0.003851	-0.004470	0.5261	0.5684
100	-0.004110	-0.004172	0.5644	0.5687
1000	-0.004127	-0.004134	0.5683	0.5687
10000	-0.004131	-0.004132	0.5687	0.5687

Table 3.6: $\int_0^1 t dW$ using 10000 paths

3.4 Stochastic differential equations

In the preface of their book [58] Kloeden and Platen state that during the past decade (the 1990s) there has been an accelerating interest in the development of numerical methods for stochastic differential equations. They add that this activity has been as strong in engineering and physical sciences as it has in mathematics. Oksendal [73] adds that SDEs have a wide range of applications

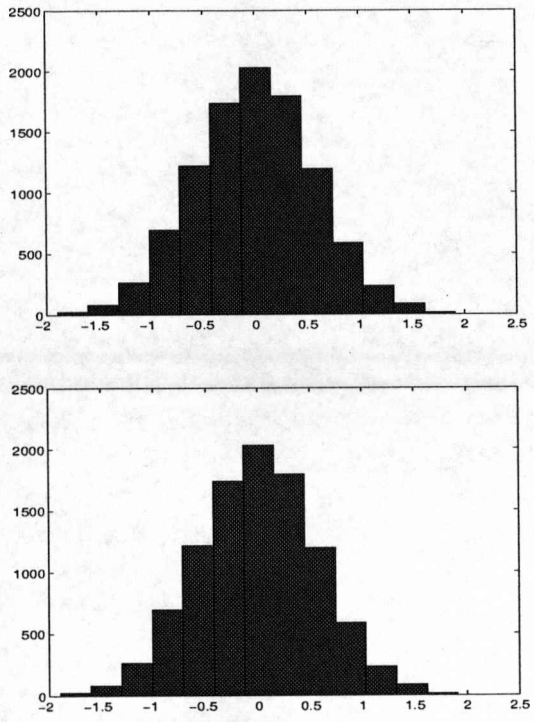


Figure 3.9: Histogram of values of $\int_0^1 tdW$ using 10000 paths
Top: *Itô* integral Bottom: Stratonovich integral

outside mathematics and that there are many fruitful connections to other mathematical disciplines and the subject has a rapidly developing life of its own as a fascinating research field with many interesting unanswered questions. As an indication of the modelling applications of SDEs we list biology [20], spread of infections [13, 19, 5], population dynamics [20] and finance [42, 52], where the Black-Scholes equation is widely used [78].

Before starting the main investigations of this thesis we need to introduce a few ideas about stochastic ordinary differential equations (SODEs) and stochastic delay differential equations (SDDEs). SODEs are defined in [73] by the equation

$$\frac{dX(t)}{dt} = f(t, X(t)) + g(t, X(t))W(t) \tag{3.1}$$

and in differential form as

$$dX(t) = f(t, X(t))dt + g(t, X(t))dW(t). \tag{3.2}$$

If $g(t, X(t)) = g(t)$, that is the function g is independent of $X(t)$, then the second term of the equation (3.2) is called *additive noise*. Otherwise this term is called

multiplicative noise, [58]. Kloeden and Platen [58] state that, unfortunately, explicitly solvable SODEs are rare in practical applications. However, to give an example of an equation which is solvable we quote the Langevin equation

$$dX(t) = -aX(t)dt + bdW(t), \tag{3.3}$$

which has solution $X(t) = X(0)\exp\left(\left(a - \frac{1}{2}b^2\right)t + bW(t)\right)$. We will use this equation in chapter 5, when we will be investigating the order of convergence of SODEs.

We can extend our definition to SDDEs by adding a delay term $X(t-\tau)$ to either or both of the functions in equation (3.2), (see [5]), to give

$$dX(t) = f(t, X(t), X(t-\tau))dt + g(t, X(t), X(t-\tau))dW(t). \tag{3.4}$$

For interest we note at this point that the software SDELab has been written at Manchester University to give solutions of SDEs using MATLAB [43].

In [5] the authors quote three different types of stability for SODEs and SDDEs, (taken from Mao [64]).

Definition (3.1 in [5]) For some $p > 0$, the null solution of the SDDE (3.4) is termed

1. locally *stable in the pth mean*, if for each $\epsilon > 0$, there exists a $\delta \geq 0$ such that $\mathbb{E}(|Y(t; t_0, \Phi)|^p) < \epsilon$ (where the notation $\mathbb{E}$ denotes the expectation) whenever $t \geq t_0$ and $\mathbb{E}(sup_{t\in[t_0-\tau,t_0]}|\Phi(t)|^p) < \delta$;

2. locally *asymptotically stable in the pth mean*, if it is stable in the pth mean and if there exists a $\delta > 0$ such that whenever $\mathbb{E}(sup_{t\in[t_0-\tau,t_0]}|\Phi(t)|^p) < \delta$ then $\mathbb{E}(|Y(t; t_0, \Phi|^p) \to 0$ for $t \to \infty$;

3. locally *exponentially stable in the pth mean* if it is stable in the pth mean and if there exists a $\delta > 0$ such that whenever $\mathbb{E}(sup_{t\in[t_0-\tau,t_0]}|\Phi(t)|^p) < \delta$ there exists some finite constant C and a $\nu^* > 0$ such that $\mathbb{E}(|Y(t; t_0, \Phi)|^p) \leq C\mathbb{E}(sup_{t\in[t_0-\tau,t_0]}|\Phi(t)|^p)exp(-\nu^*(t-t_0))$ for $t_0 \leq t \leq \infty$.

If, in the above, δ may be taken as arbitrarily large then the stability is in each case *global* rather than local [5]. The authors of [5] go on to illustrate a different approach to stability, that of *stochastic stability* or *stability in probability*, with a parallel to one of the above definitions:

Definition (3.2 in [5]) The null solution of the SDDE (3.4) is termed *stochastically stable in probability*, if for each $e \in (0,1)$ and $\epsilon > 0$, there exists a $\delta \equiv \delta(e, \epsilon) \geq 0$ such that $P(|Y(t; t_0, \Phi)| \leq \epsilon$ for all $t \geq t_0) \geq 1 - e$ (where P denotes probability) whenever $t \geq t_0$ and $sup_{t\in[t_0-\tau,t_0]}|\Phi(t)|^p < \delta$ with probability 1. Finally, the authors add that extensions for the definitions of stability to apply to non-null solutions follow the usual lines. Stability of the stochastic logistic model is discussed in [44], while Burrage et al give an overview in [17].

As mentioned earlier, Kloeden and Platen stated that explicitly solvable SODEs are rare. Once we add the delay to the problem this statement becomes much stronger and it is very unlikely that we will find an explicit solution to a SDDE. To reach any conclusion with these equations we must resort to experimental or Monte Carlo techniques, and this is the route followed for our investigations throughout the remainder of this thesis.

Chapter 4

Early results: setting the objectives

[1] The general form of stochastic delay differential equation that we would like to consider takes the form

$$Y(t) = Y(t_0) + \int_{t_0}^{t} F(s, Y(s), Y(s-\tau))ds + \int_{t_0}^{t} G(s, Y(s), Y(s-\tau))dW(s) \tag{4.1}$$

with $Y(t) = \Phi(t)$ for $t \in [t_0 - \tau, t_0]$.

This equation is often written, in the Itô sense, in the shorthand form

$$\begin{aligned} dY(t) &= F(t, Y(t), Y(t-\tau))dt + G(t, Y(t), Y(t-\tau))dW(t), \qquad t \geq t_0 \\ Y(t) &= \Phi(t), \qquad t \in [t_0 - \tau, t_0] \end{aligned} \tag{4.2}$$

where τ is the constant *time-lag* and $W(t)$ is a standard *Wiener* process. Following the terminology used in [5], F is called the *drift* term and G is the *diffusion* term.

The analysis of equations of the general form (4.1) is still under development and there is comparatively little known about the qualitative behaviour of solutions of such a general equation as $t \to \infty$. For this reason, we choose to restrict our attention to a simple linear test equation. Despite its simplicity, the test equation continues to present challenges both to classical and numerical analysis. In our test equations we take the time-lag $\tau = 1$.

$$\begin{aligned} dY(t) &= \lambda Y(t-1)dt + \mu Y(t)dW(t), \qquad t \geq 0 \\ Y(t) &= t + \frac{1}{2}, \qquad t \in [-1, 0]. \end{aligned} \tag{4.3}$$

[1] The content and the approach in this chapter appeared as 'Predicting changes in dynamical behaviour in solutions to stochastic delay differential equations' in Communications on Pure and Applied Analysis, Volume 5, Number 2, June 2006, pp. 367-382, [71].

We shall also refer to the stochastic delay logistic equation of the form

$$\begin{aligned} dY(t) &= \lambda Y(t-1)[1+Y(t)]dt + \mu Y(t)dW(t), \qquad t \geq 0 \\ Y(t) &= t + \frac{1}{2}, \qquad t \in [-1,0]. \end{aligned} \tag{4.4}$$

$Y(t) = t + \frac{1}{2}$ was selected as a simple bounded non-constant initial function for our test analyses.

Equation (4.3) arises when one linearises equation (4.4) about the zero solution. Classically one uses results about the deterministic linear equation ($\mu = 0$ in (4.3))

$$\begin{aligned} dY(t) &= \lambda Y(t-1)dt, \qquad t \geq 0 \\ Y(t) &= t + \frac{1}{2}, \qquad t \in [-1,0]. \end{aligned} \tag{4.5}$$

to make predictions about the stability of solutions to the delay logistic equation ($\mu = 0$ in (4.4))

$$\begin{aligned} dY(t) &= \lambda Y(t-1)[1+Y(t)]dt, \qquad t \geq 0 \\ Y(t) &= t + \frac{1}{2}, \qquad t \in [-1,0]. \end{aligned} \tag{4.6}$$

Of particular interest to us is the investigation of the analogous behaviour (in the stochastic case) of the loss of stability in the deterministic linear case and the Hopf bifurcation in the deterministic logistic equation.

The equations (4.5,4.6) are known to have a bifurcation, where there is a fundamental change in the qualitative behaviour of solutions, at the value $\lambda = -\frac{\pi}{2}$ (for example, see [70] p.17–19, or [45]). Figure 4.1 illustrates this change in behaviour for the linear equation (4.5).

Remark 4.0.1 Naturally one can express equations (4.5) and (4.6) in the form of an integrodifferential equation with a singular kernel. In this way, delay differential equations may be used to give more general insight into analytical methods for other classes of functional differential equation ([24, 48]).

4.1 Basic numerical methods

Numerical methods for stochastic differential equations with delays are very under-developed. They must usually be adapted (with considerable care) from methods either for deterministic delay differential equations, or for stochastic ordinary differential equations. Direct analysis of methods for SDDEs has been considered, for example, in [6, 8, 7, 10, 54, 55]. We refer the reader also to the

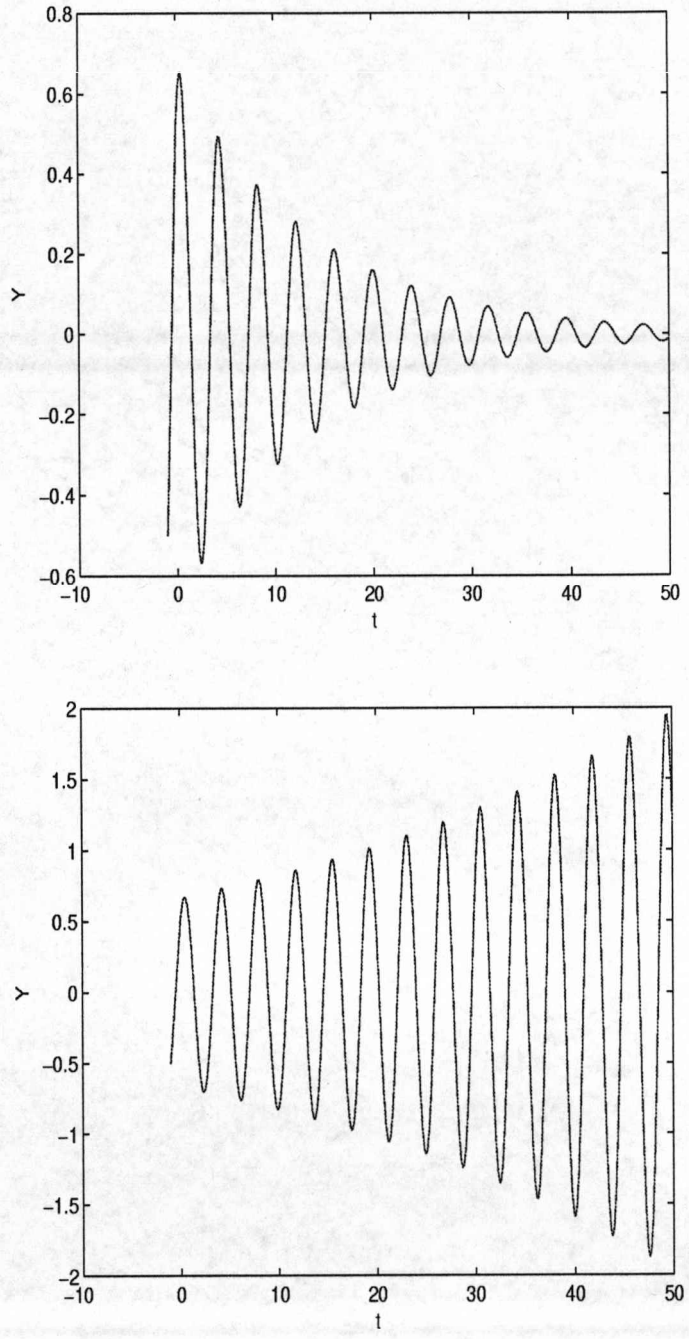


Figure 4.1: Forward Euler, step size = 0.1, applied to equation (4.5).
Top: $\lambda = -1.5 > -\frac{\pi}{2}$
Bottom: $\lambda = -1.7 < -\frac{\pi}{2}$

works [17, 50, 51, 58] for background material for stochastic ordinary differential equations and to the recent book [12] for a survey of methods for delay differential equations.

In this chapter we base our discussions on the standard linear ϑ-methods for the deterministic equations. We take $\vartheta = 0, 0.5, 1$, corresponding to the forward Euler, trapezium and backward Euler methods. We also use the third order implicit Adams Moulton method.

For the deterministic equation

$$\begin{aligned} Y'(t) &= F(t, Y(t), Y(t-\tau)), \qquad t \geq t_0 \\ Y(t) &= \Phi(t), \qquad t \in [t_0 - \tau, t_0] \end{aligned} \tag{4.7}$$

the ϑ methods, for fixed $h > 0$, take the form

$$y_{n+1} = y_n + (1-\vartheta)hF_n + \vartheta h F_{n+1},$$

and the Adams Moulton method takes the form

$$y_{n+1} = y_n + \frac{h}{12}(5F_{n+1} + 8F_n - F_{n-1})$$

where $Nh = \tau$, $y_n \simeq Y(nh)$, $F_n = F(nh, y_n, y_{n-N})$ and $y_{-N}, \ldots, y_0$ are given by our initial function.

For the stochastic equations we use the stochastic variant of the ϑ-methods, often known as the semi-implicit Euler methods (see [50]). We again use $\vartheta = 0, 0.5, 1$. $\vartheta = 0$ gives the classical Euler-Maruyama method. These schemes are described more fully in section 5.2.

4.2 A phenomenological approach to bifurcations

4.2.1 A deterministic equation

The investigation of bifurcation problems using numerical methods is well established. The work of [25, 26, 46] concentrates on ordinary differential equations. For delay differential equations, we refer to the works [30, 31, 32, 63] as well as [81]. For stochastic delay differential equations there has been little previous work published (see [5]).

We describe here the results of simple numerical experiments that were carried out on the deterministic equations in order to get an insight into the bifurcation problem.The idea of the phenomenological approach is that one should be able to detect *by eye* changes in the dynamical behaviour of the solutions.

With this in mind, we present numerical solutions of the linear equation (4.5) computed over the interval [0, 200000] using a variety of step sizes with a range of λ values close to $-\frac{\pi}{2}$ which is the analytical bifurcation value for this equation shown earlier. For each method and step size we focused on the value of $\lambda = \lambda_{bif}$, where the graph of the numerical solution appeared to change. Figure 4.2 indicates the way the solutions change over this time interval.

We seek to find λ_{bif}, the value of λ where the behaviour changes. Figure 4.2 shows that for $\lambda = -1.4945$ the solution, using forward Euler with $h = 0.1$, has converged to $Y(t) = 0$ as $t \to \infty$, whereas for $\lambda = -1.4947$ the solution is oscillating with a rapidly increasing amplitude, and $|Y(t)|$ is unbounded as $t \to \infty$. The figure shows $\lambda_{bif} \in [-1.4947, -1.4945]$ for the forward Euler method with a step size of 0.1. By producing further diagrams with increasing lower limit and decreasing upper limit it was possible to obtain the point at which the shape changed with a some accuracy. Figure 4.3 shows the method approaching the value of $\lambda = -1.494602$ for the forward Euler method with $h = 0.1$.

We repeated this technique for step sizes of 0.2, 0.1, 0.025 and also used the backward Euler, trapezium and implicit Adams Moulton methods. The observed values of λ_{bif} for each step-size are shown in Table 4.1.

h	Forward Euler	Backward Euler	Trapezium	Implicit Adams Moulton
0.2	-1.423148	-1.736482	-1.5838444	-1.56974848
0.1	-1.494602	-1.651587	-1.5740341	-1.57064978
0.05	-1.532109	-1.610638	-1.5716043	-1.57077707
0.025	-1.551306	-1.590575	-1.5709982	-1.57079386

Table 4.1: P-bifurcation values, λ_{bif} (Analytical value $= -\frac{\pi}{2} = -1.57079633$)

Since we know the exact values $-\frac{\pi}{2}$ at which the true bifurcation happens, we can calculate the errors in the apparent bifurcation values. These are given in Table 4.2.

One can immediately see that the changes in behaviour are detected at different parameter values, according to the choice of method, and the step length h. We can investigate the relationship more precisely: By defining $d = \left|\lambda_{bif} - (-\frac{\pi}{2})\right|$ we can now plot the graphs of $\log(d)$ against $\log(h)$ for each method, and add the least squares regression lines to our plots, see [77]. All four lines have correlation coefficients of $r = 1$, confirming the excellent fits [77]. The regression analysis gives us the following approximations for the bifurcation values of λ in terms of h,

$\lambda_{bif} \approx -1.568802 + 0.730786h$ for forward Euler;

$\lambda_{bif} \approx -1.569027 - 0.835127h$ for backward Euler;

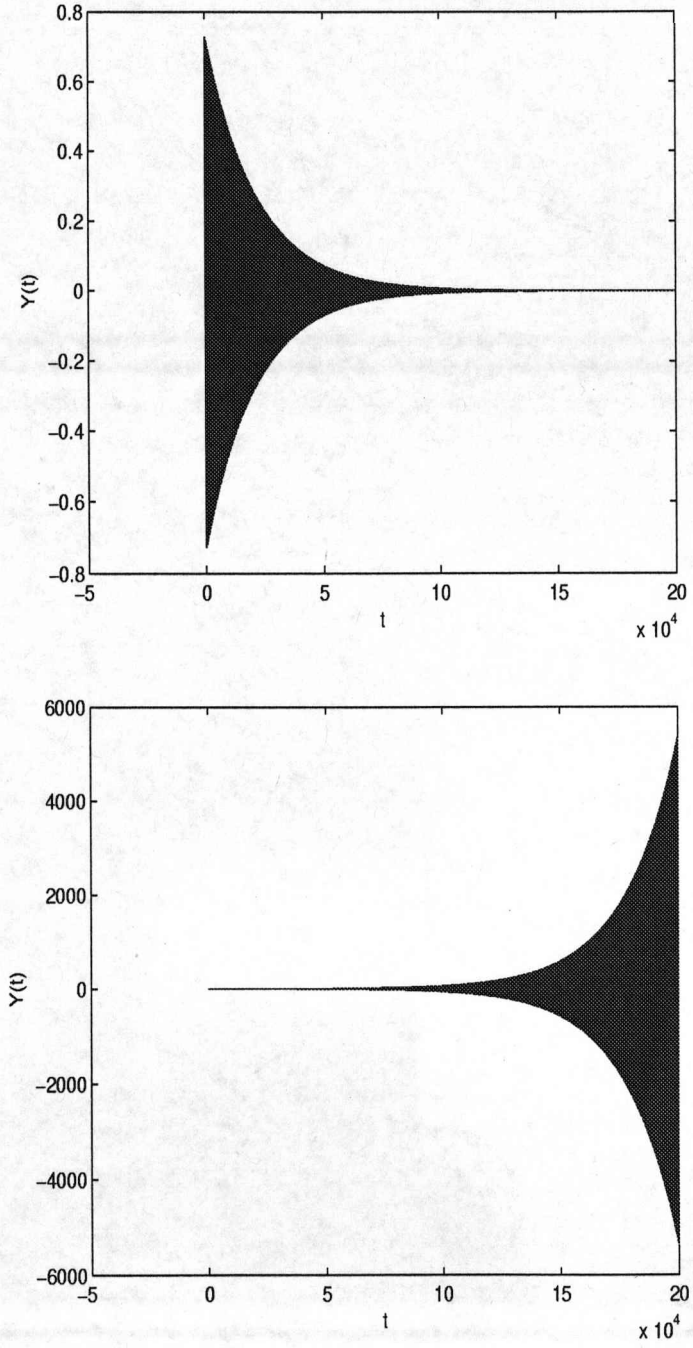


Figure 4.2: Forward Euler, step size h = 0.1
Top: $\lambda = -1.4945 > \lambda_{bif}$
Bottom: $\lambda = -1.4947 < \lambda_{bif}$

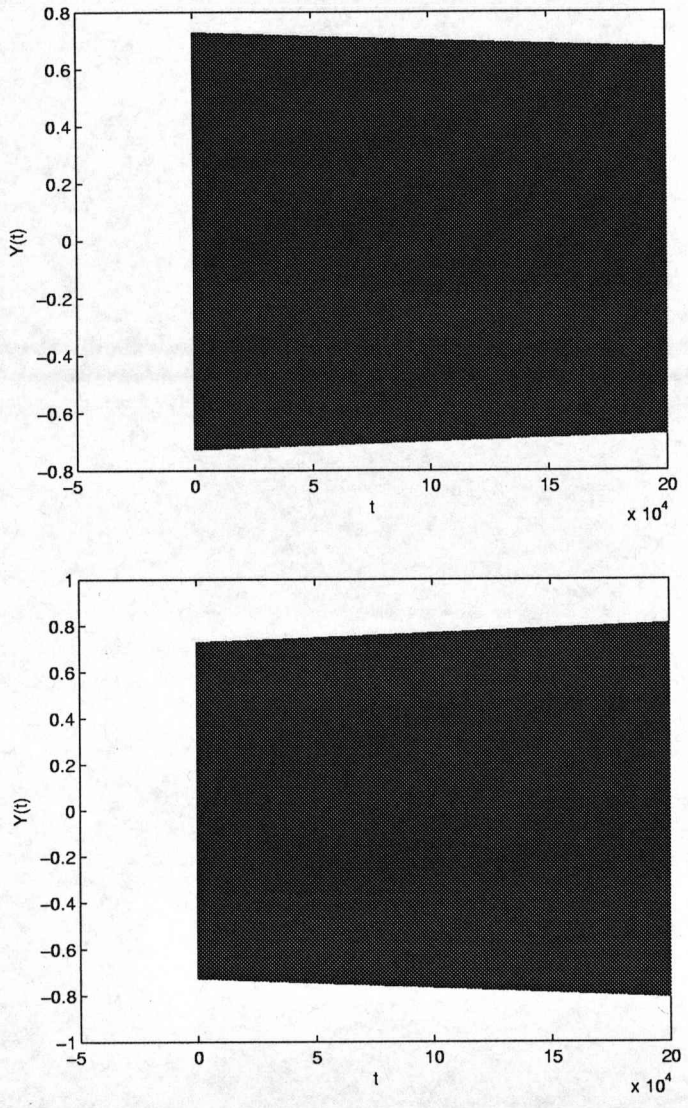


Figure 4.3: Forward Euler, step size h = 0.1
Top: $\lambda = -1.494601 > \lambda_{bif}$
Bottom: $\lambda = -1.494603 < \lambda_{bif}$

$\lambda_{bif} \approx -1.570785 - 0.326388h^2$ for the trapezium rule;
$\lambda_{bif} \approx -1.570790 + 0.130364h^3$ for the implicit Adams Moulton rule.

The plots are shown in figure 4.4 and gradients of these lines are given in table 4.3.

The results of these simple experiments indicate that the numerical method advances or retards the apparent bifurcation value of λ, depending upon the method, and the error in the approximation of the exact bifurcation point varies

h	Forward Euler	Backward Euler	Trapezium	Implicit Adams Moulton
0.2	0.147648	-0.165686	-0.0130481	0.00104785
0.1	0.076194	-0.080790	-0.0032378	0.00014655
0.05	0.038687	-0.039842	-0.0008080	0.00001926
0.025	0.019490	-0.019779	-0.0002018	0.00000247

Table 4.2: Error, $\lambda_{bif} - (-\frac{\pi}{2})$.

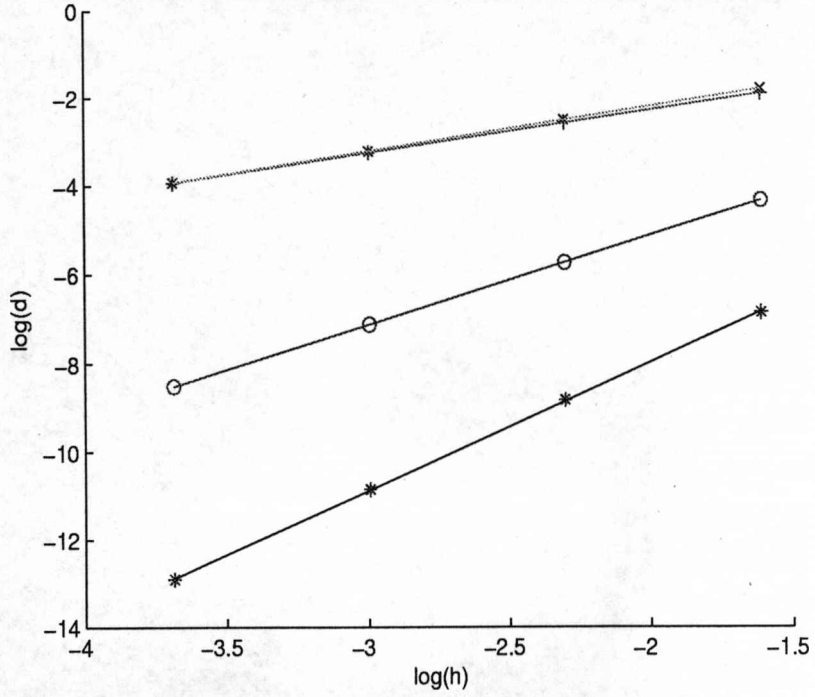


Figure 4.4: Variation of perturbation with step size
+ Forward Euler x Backward Euler,
o Trapezium * Implicit Adams Moulton.

as h^n, where n is the order of the method. If we let h tend to zero we can obtain the bifurcation values for each numerical scheme predicted by our technique. Table 4.4 shows these values, which are all close to the theoretical value of $-\frac{\pi}{2} = -1.57079633$, with the this comparison improving as the order of the numerical scheme increases.

In a sense, this should come as no surprise. In the works [37, 38, 39, 40,

Method	Gradient	Order of Method
Forward Euler	0.974	1
Backward Euler	1.022	1
Trapezium	2.005	2
Adams Moulton	2.911	3

Table 4.3: Gradients of the log graphs

Method	λ_{bif}
Forward Euler	-1.568802
Backward Euler	-1.569027
Trapezium	-1.570785
Adams Moulton	-1.570790

Table 4.4: Phenomenological estimates of bifurcation value

81] it was shown that suitable numerical methods applied close to this type of bifurcation induce discrete schemes that display the same type of change in behaviour. Indeed, it was shown that the approximation of the parameter value where the change in behaviour occurs is to the order of the method being used. The point of this experiment is that we have shown that one can detect this change of behaviour quite accurately by eye, and this validates the basic idea of a phenomenological approach to finding bifurcations that we shall apply to the stochastic equation in the following section.

However, although the analytical results described in the previous paragraph apply equally to certain non-linear equations such as the delay logistic equation, the phenomenological approach is not so effective now. The logistic equation undergoes a Hopf bifurcation at $\lambda = -\frac{\pi}{2}$ and $Y(t)$ may *oscillate* for parameter values less than $-\frac{\pi}{2}$ instead of becoming unbounded, as in the linear case. Figure 4.5 shows the solution, using forward Euler with stepsize $h = 0.1$, appearing to converge for $\lambda = -1.4945$ and to oscillate for $\lambda = -1.4947$. It is much more difficult to detect the differences between persistent oscillations and very slowly growing or decaying oscillations over any finite time interval and therefore there is considerable scope for uncertainty in predictions of the parameter value where solution behaviour changes.

We can obtain more insight into the logistic bifurcation by considering the *amplitude* of the oscillatory solutions. The variation of $\sup_{[199900,200000]}(|Y(t)|)$ with λ is shown in figure 4.6. This figure suggests that $\lambda_{bif} \approx -1.49454$. It is not surprising that this value is a reasonable approximation to the value -1.494602 obtained from the linear analysis.

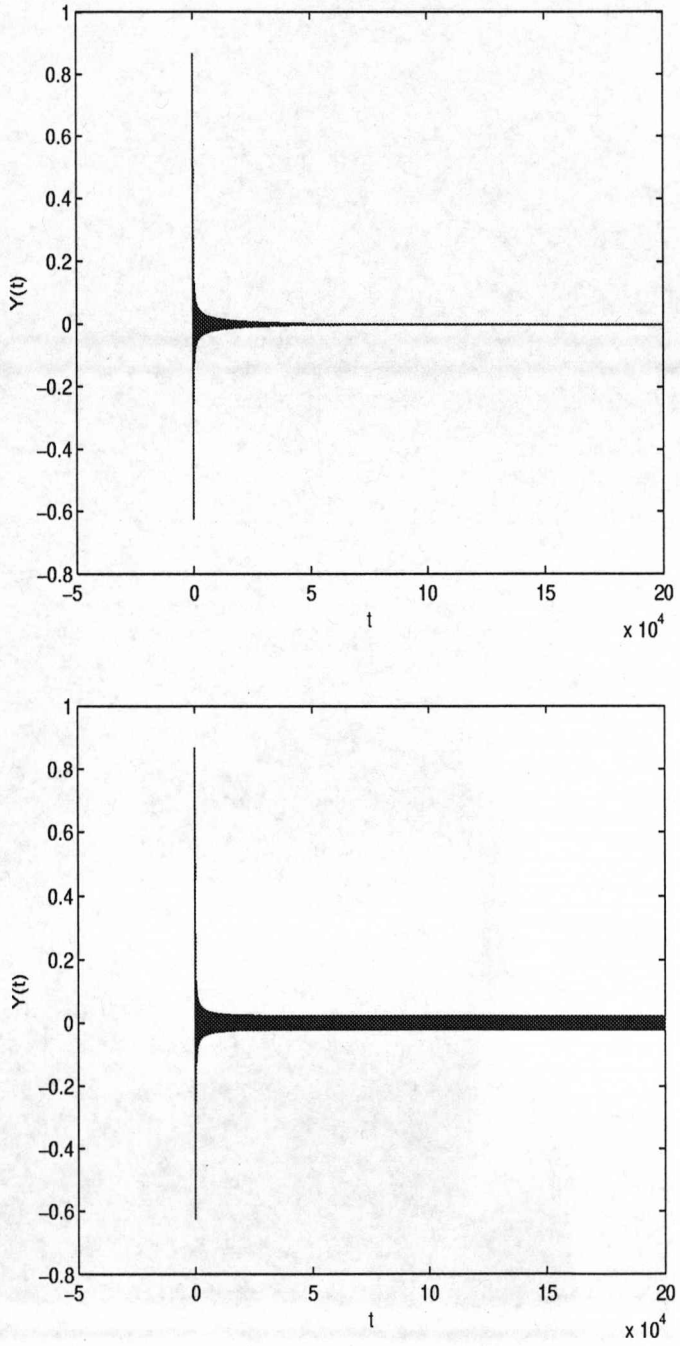


Figure 4.5: Logistic equation, forward Euler, step size = 0.1
Top: $\lambda = -1.4945 > \lambda_{bif}$ Bottom: $\lambda = -1.4947 < \lambda_{bif}$?

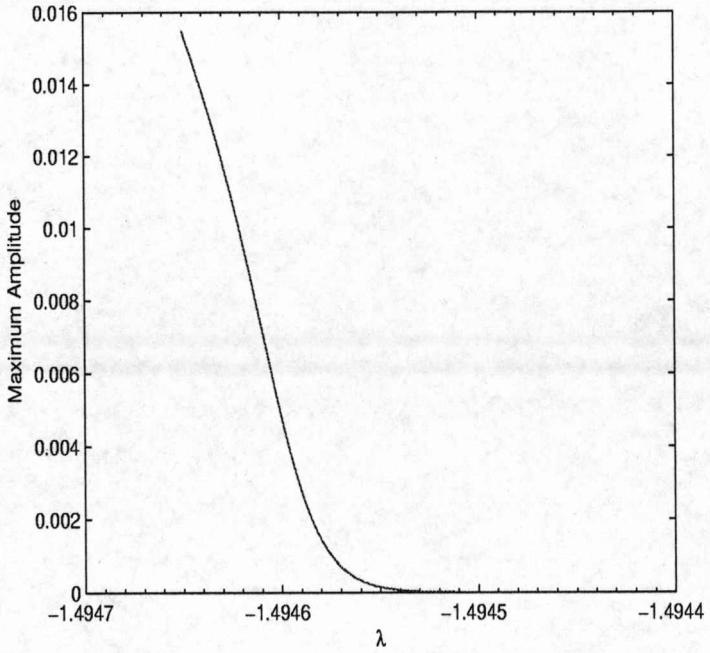


Figure 4.6: Logistic equation, forward Euler, step size = 0.1

4.2.2 Application to a stochastic equation

We carried out similar numerical experiments in order to gain some insight into the behaviour of the solution to equation (4.3), the stochastic variant of equation (4.5), for values of λ close to $-\frac{\pi}{2}$, the bifurcation value for the deterministic equation. For each value of λ we used the Euler-Maruyama method

$$y_{n+1} = y_n + h\lambda y_{n-N} + \mu y_n \Delta W_n \tag{4.8}$$

where $Nh = 1$ and $y_{-N}, \ldots, y_0$ are given by our initial function, to obtain a numerical solution in the interval [0, 200000]. Of course, when $\mu = 0$ this method reduces to the forward Euler method for the deterministic equation. We could present results for other *stochastic $\vartheta-$methods* but remark that no fresh insight comes from these alternative schemes.

As shown in chapter 3, using Matlab's random number generator it is possible to simulate the values of $dW(t)$ to repeat an identical Brownian motion path. Using this single path will ensure that only λ is varied in successive trajectories. Taking $\mu = 0.1$ and maintaining the stepsize $h = 0.1$ for consistency with the deterministic case we used the Euler-Maruyama method to obtain solutions to equation (4.3) for values of λ close to the bifurcation value. Figure 4.7 shows six solutions.

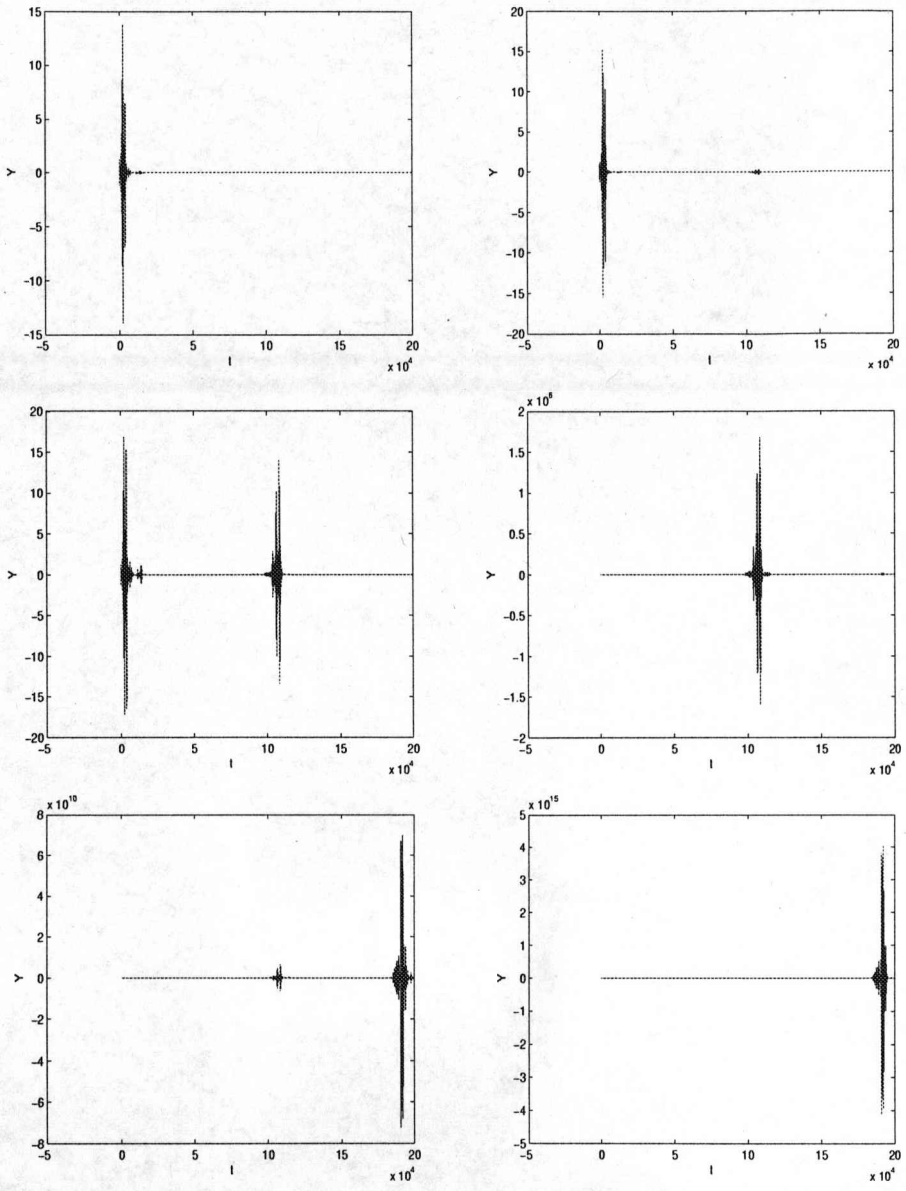


Figure 4.7: Trajectories of equation (4.3) for $\mu = 0.1$ and stepsize $h = 0.1$, using the fixed initial Brownian path in every trajectory.

Left:	Right:
$\lambda = -1.4925$	$\lambda = -1.4926$
$\lambda = -1.4927$	$\lambda = -1.4928$
$\lambda = -1.4929$	$\lambda = -1.4930$

As we can see from figure 4.7 the value of the parameter λ at which the solution changes behaviour is not as clearly defined as in the deterministic case.

As λ increases from -1.4925 to -1.4930 we can see that for the first trajectory $|Y(t)| \to 0$ very quickly as $t \to \infty$. The second trajectory has a small interval of small oscillations at $t \approx 110000$, but again $|Y(t)| \to 0$ as $t \to \infty$. This oscillation has a much larger amplitude during the third trajectory. This amplitude is greatly increased in the fourth trajectory and a further period of oscillation is evident as t approaches 200000. The final two trajectories show $|Y(t)|$ is unbounded. Therefore one cannot tell from a single trajectory a precise parameter value at which all trajectories change in behaviour.

We can also find $\sup|Y(t)|$ for $t \in [190000, 200000]$, and this value was plotted in figure 4.8 for values of $\lambda \approx -1.4972$. This figure also shows that we can only *estimate* an interval of values for λ_{bif}, with $\sup|Y(t)|$ approaching zero for a number of values of λ between -1.492745 and -1.492695.

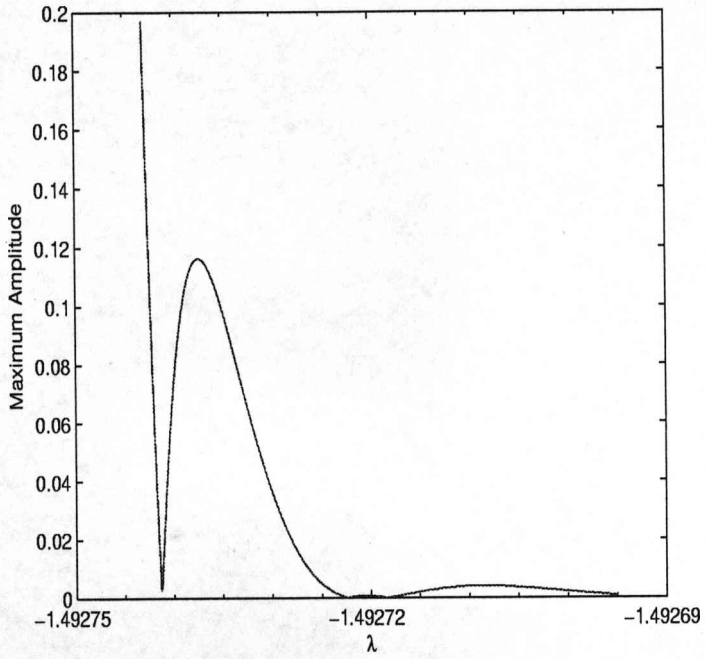


Figure 4.8: Variation with λ of the maximum amplitude of $Y(t)$ for $t \in [190000, 200000]$.

However, for the stochastic equation, it is meaningless to talk about *the solution* because each solution depends on a particular Brownian motion path. Figure 4.9 shows that if we use a different Brownian motion path then the interval containing λ_{bif} (in this case $(-1.4935, -1.4930)$) is disjoint from the interval determined using the first Brownian path (which was $(-1.4930, -1.4925)$).

At this stage in the investigation we have used a single coefficient of noise, $\mu = 0.1$. Our bifurcation range with $\mu = 0.1$ does not include the value from the

deterministic case $\mu = 0$, so it would be reasonable to expect that a different noise coefficient will move the interval of bifurcation. Figure 4.10 shows trajectories with $\mu = 0.3$ for the same Brownian path that produced figure 4.9. For $\mu = 0.3$ the behaviour of the solution is now in the region of $(-1.4805, -1.4795)$.

Therefore we need to consider the simulation of solution trajectories. To take account of the fact that different Brownian motion paths might, in principle at least, lead to quite different qualitative behaviour of the solution, it will be necessary to consider a large number of solution trajectories for each choice of method, parameter values and step length so that we can see whether common features can be deduced.

The phenomenological approach of estimating the value of λ_{bif} *by eye* will be inadequate to analyse such a large number of trajectories.

Considering the results of our basic experiments in this chapter we come up with the following objectives for this project:

1. find a better way of identifying the changes in behaviour for the stochastic case,
2. forecast the bifurcation value for each $\vartheta - method$,
3. forecast the bifurcation value as h varies,
4. forecast the bifurcation value as μ varies.

Finally, we can test the conjecture that the deterministic case will be extended to the stochastic case for objectives 2 and 3.

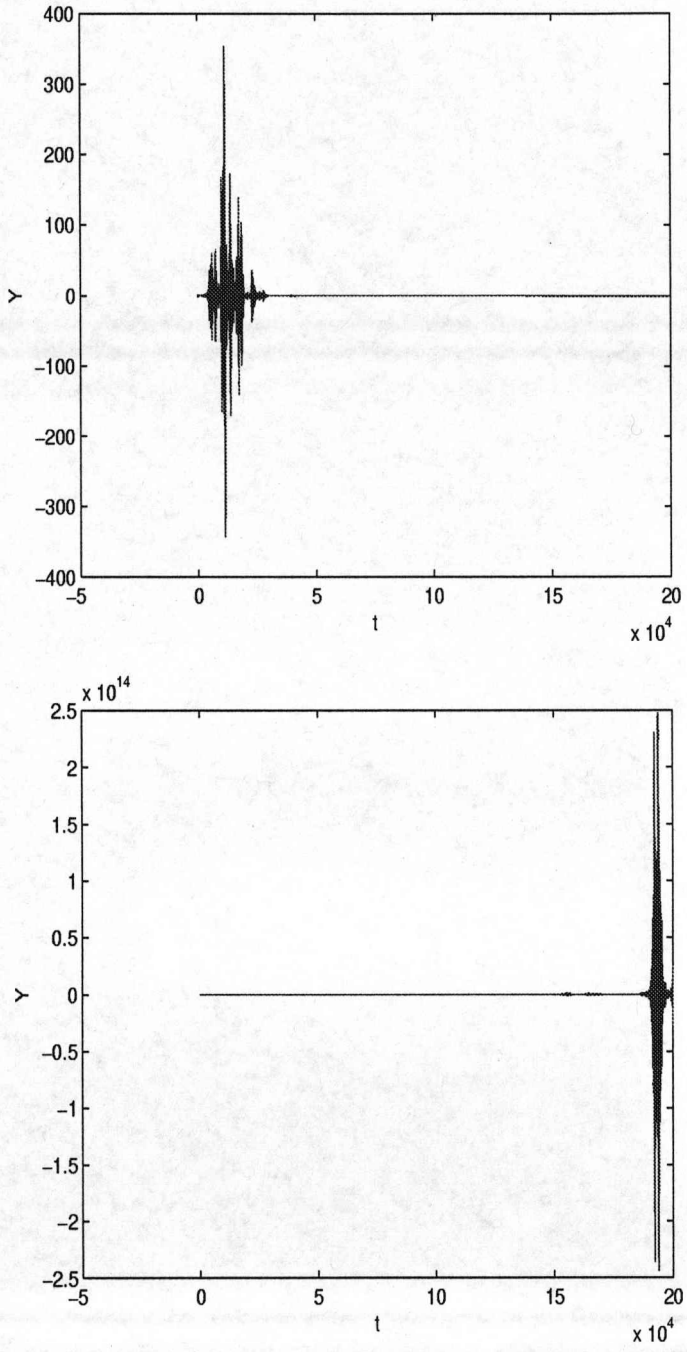


Figure 4.9: Trajectories of equation (4.3) for $\mu = 0.1$ and stepsize $h = 0.1$, using a second fixed Brownian path.

Left: $\lambda = -1.4930$

Right: $\lambda = -1.4935$

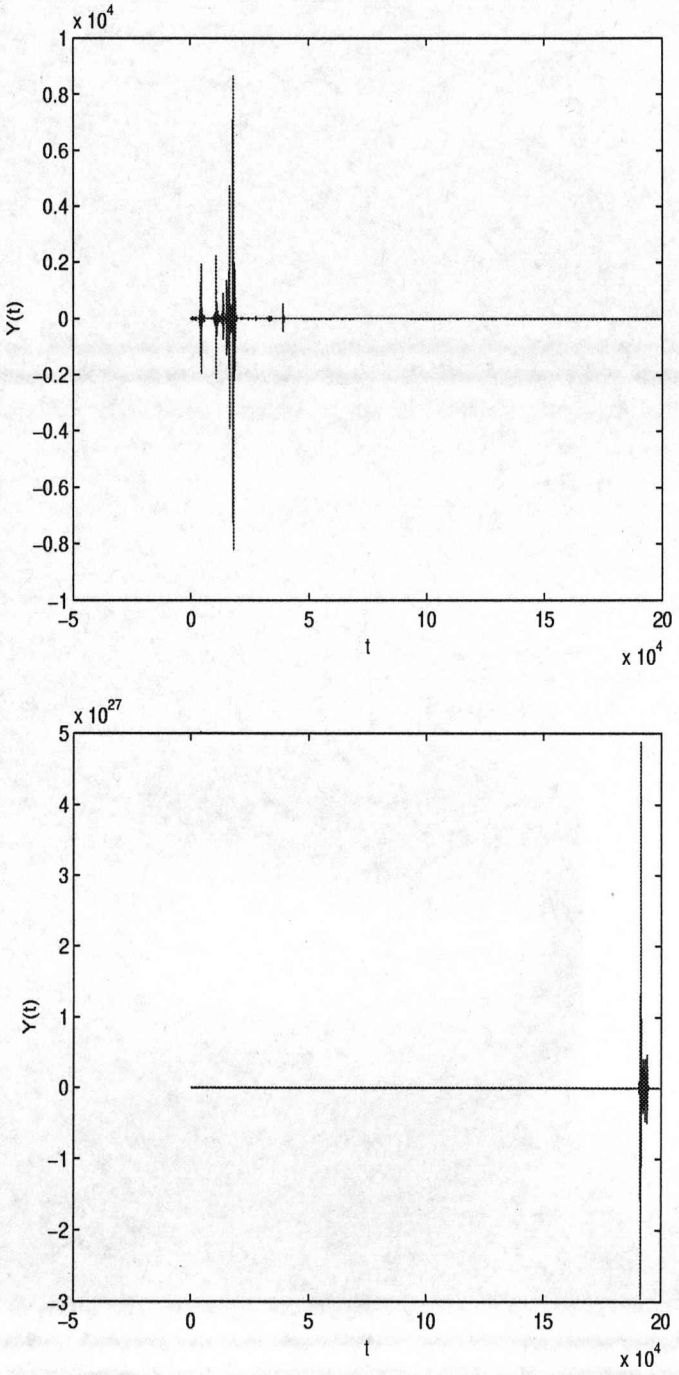


Figure 4.10: Trajectories of equation (4.3) for $\mu = 0.3$ and stepsize $h = 0.1$, applying the fixed Brownian path used for figure 4.9.
Left: $\lambda = -1.4795$ Right: $\lambda = -1.4805$

Chapter 5

Towards greater precision: P-bifurcations and D-bifurcations

[2] In the previous chapter we considered the use of the Euler-Maruyama method for detecting bifurcation points in the numerical solution of (4.3). We adopted a phenomenological approach: we sought changes in the dynamical behaviour of solutions to the equation. In [5] the authors produced the following summarised findings:

1. for equation (4.3) they detected changes in the dynamical behaviour of solution trajectories as λ varies close to $-\pi/2$
2. the value of λ at which the behaviour changes varies with the step length h and the coefficient of noise
3. it is quite hard to detect the precise parameter value at which changes have occurred
4. they identified *intervals* for the parameter value λ over which the solution trajectories changed their behaviour
5. they conjectured that there may be no precise parameter value at which *every* approximate trajectory changes its behaviour.

In the context of known theory for stochastic functional differential equations (see, for example, [68, 69]) these observations are somewhat surprising. One usually hopes to see the underlying dynamical behaviour of all the solution trajectories to change in the same way as the parameter λ passes through some critical value. The fact that this property does not show up in the investigations in [5] could be caused by one of the following:

[2]The content and the approach in this chapter appeared as 'Numerical Investigation of D-bifurcations for a stochastic delay logistic equation' in Stochastics and Dynamics, Volume 5, Number 2 (2005), pp. 211-222, [34].

1. it could be a property of the numerical method used (or indeed of all numerical methods) that the underlying dynamics are not recreated faithfully in the discrete problem

2. it could be a feature of the phenomenological approach to detecting changes in behaviour. The phenomenological approach requires interpretation of changes in dynamical behaviour of solution trajectories and a more precise approach would be useful

3. it could be the result of the fact that the theory given in [68, 69] describes asymptotic results as $t \to \infty$. Clearly this situation is not easily replicated in numerical experiments since we must derive our results for some (large) finite time.

The methods adopted in the current chapter are designed to address these uncertainties.

5.1 The attraction of a D-bifurcation approach

The phenomenological approach adopted in previous work concerns the detection of subtle changes in the solution trajectories as the parameter λ varies. The approach leads to the identification of so-called P-bifurcations. An alternative approach is to look for changes in the underlying dynamical system through calculating values of Lyapunov exponents and detecting parameter values at which a Lyapunov exponent changes sign. This leads to the so-called D-bifurcation method.

We do not give a full discussion of Lyapunov exponents in this thesis and for more details we refer, for example, to [1, 2]. To define a Lyapunov exponent we need to have a linear equation (or the linearised form of a nonlinear problem), and we use a particular moment (in this case the first moment). A linear stochastic delay equation has infinitely many Lyapunov exponents (see [5]) and we are interested here in sign changes of the principal (right-most) Lyapunov exponent in the complex plane. It turns out that this value, Λ, is simple to define in terms of expectations of the solution:

$$\Lambda = \lim_{t \to \infty} \sup E(\frac{1}{t} \log |Y(t)|) \tag{5.1}$$

In this thesis we consider the approximation of Lyapunov exponents through solution trajectories simulated by numerical methods applied to equations (4.3,4.4). Therefore we are undertaking a form of numerical D-bifurcation analysis, [1].

5.2 The numerical methods under consideration

Previous results for the deterministic equation (4.5) under numerical approximation motivates us to consider a family of numerical schemes for the stochastic equation (4.3). This approach will help us to understand how changing the numerical method influences the conclusions we might draw about bifurcations. One such family for deterministic equations is the set of linear ϑ methods. Here we shall use their stochastic counterparts. Stochastic $\vartheta-$methods for stochastic ordinary differential equations are analysed in [50]. With these methods any implicitness of the method is applied only to the drift term and this avoids the known instability issues when implicit methods are applied directly to stochastic differential equations. Thus the deterministic $\vartheta-$method for solving

$$y'(t) = \lambda y(t); \quad y(0) = \alpha \tag{5.2}$$

takes the form

$$y_{n+1} = y_n + (1-\vartheta)h\lambda y_n + \vartheta h\lambda y_{n+1}; \quad y_0 = \alpha \tag{5.3}$$

and is adapted for the stochastic problem

$$dY(t) = \lambda Y(t)dt + \mu Y(t)dW(t); \quad y(0) = \alpha \tag{5.4}$$

to take the form

$$Y_{n+1} = Y_n + (1-\vartheta)h\lambda Y_n + \vartheta h\lambda Y_{n+1} + \mu Y_n \triangle W_n; \quad Y_0 = \alpha. \tag{5.5}$$

Methods of this type can be called semi-implicit Euler methods. Of course, when $\vartheta = 0$ the method reduces to the classical Euler-Maruyama method.

These stochastic $\vartheta-$methods can be adapted quite naturally for stochastic delay differential equations. Thus, we shall use the schemes

$$Y_{n+1} = Y_n + (1-\vartheta)h\lambda Y_{n-N} + \vartheta h\lambda Y_{n+1-N} + \mu Y_n \triangle W_n \tag{5.6}$$

and

$$Y_{n+1} = Y_n + (1-\vartheta)h\lambda Y_{n-N}(1+Y_n) + \vartheta h\lambda Y_{n+1-N}(1+Y_{n+1}) + \mu Y_n \triangle W_n \tag{5.7}$$

(where $Nh = 1$ and $Y_{-N}, ..., Y_0$ are given by the initial function) for approximating the solution trajectories of equations (4.3,4.4) respectively. We shall concentrate on the most common choices of $\vartheta = 0, \frac{1}{2}, 1$.

5.3 Convergence

Before we use the three schemes described in the previous section we need to determine if the schemes are all reliable. In [49] two types of convergence are defined and investigated for the Euler-Maruyama ($\vartheta = 0$) method on SDEs. We will extend this work to also include the $\vartheta = 0.5$ and $\vartheta = 1$ methods. We need to test our methods on an equation with a known solution. For this reason we will apply our methods to a SDE rather than a SDDE[3]. It is known that (see [58]) the Black-Scholes equation

$$dY(t) = \lambda Y(t)dt + \mu Y(t)dW(t), \qquad Y(0) = Y_0 \tag{5.8}$$

has the solution

$$Y(t) = Y_0 e^{(\lambda - 0.5\mu^2)t + \mu W(t)} \tag{5.9}$$

We state the definitions given in [49]:
Strong convergence:
A method is said to have strong order of convergence γ if there exists a constant C such that, for sufficiently small step length h, $E|Y_n - Y(t_n)| \leq Ch^\gamma$ where $Y(t_n)$ is the exact value of the solution at gridpoint $t_n \in [0, 1]$.
Weak convergence:
A method is said to have weak order of convergence γ if there exists a constant C such that, for sufficiently small step length h, $|E(Y_n) - E(Y(t_n))| \leq Ch^\gamma$.
In [49] Higham shows that the Euler-Maruyama method has strong order of convergence $\gamma = 0.5$ and weak order of convergence $\gamma = 1$. It is noted that if $\mu = 0$, giving us a deterministic equation, the strong order of convergence is 1. We will now adapt the method of Higham to produce the orders of convergence for all three of our ϑ-methods.

5.3.1 Strong convergence

Basically, this is concerned with the expected error at each gridpoint. We plan to find the solution over the interval $[0, 1]$ and to find the expected error at the endpoint, $t = 1$, which we would expect to have the maximum error over the interval. The SDE we will use is

$$dY(t) = 3Y(t) + 2Y(t)dW(t), \qquad Y(0) = 1$$

which has the solution

$$Y(t) = e^{t + 2W(t)}$$

[3]We realise here that the convergence of a numerical method for a SDE does not guarantee its convergence for a SDDE, just as we have already observed that a scheme converging for an ODE doesn't guarantee its convergence for a DDE. However, with no analytical solutions available to test on SDDEs, this section was considered to be a compromise.

and hence for our experiments the exact solution at $t = 1$ is

$$Y(1) = e^{1+2W(1)}$$

Our MATLAB code produces 5000 discrete Brownian paths over $[0, 1]$ with $\delta t = \frac{1}{4000}$. We use each path with the 9 step lengths $h = \frac{1}{4000}, \frac{1}{2000}, \frac{1}{1000}, \frac{1}{800}, \frac{1}{500}, \frac{1}{400}, \frac{1}{250}, \frac{1}{200}, \frac{1}{100}$ to estimate Y_n, our approximation to $Y(1)$ for every path and for each ϑ. For each step length we calculate the 5000 values of the absolute error $|Y_N - Y(1)|$ and estimate $E|Y_n - Y(t_n)|$ by the mean of these 5000 values.

If we write

$$E_S = |Y_n - Y(t_n)|$$

then the definition of strong convergence requires

$$E_S \leq Ch^\gamma$$

or, by taking logarithms

$$\log(E_S) \leq \gamma \log(h) + \log(C).$$

We can plot the nine points of $\log(E_S)$ against $\log(h)$ for each method and these are shown in figures 5.1, 5.2 and 5.3. We can see excellent linear relationships with all three plots, so we have added the least squares regression lines.
For $\vartheta = 0$ the equation of the regression line is

$$\log(E_S) \approx 0.5224 \log(h) + 1.8780$$

which gives us

$$E_S \approx 6.5404 h^{0.5224}.$$

For $\vartheta = 0.5$ the equation of the regression line is

$$\log(E_S) \approx 0.5289 \log(h) + 1.8999$$

which gives us

$$E_S \approx 6.6852 h^{0.5289}.$$

Finally for $\vartheta = 1$ the equation of the regression line is

$$\log(E_S) \approx 0.5396 \log(h) + 1.9395$$

which gives us

$$E_S \approx 6.9553 h^{0.5396}.$$

All three cases imply that all our methods have a strong order of convergence $\gamma = 0.5$. We repeat here that for the deterministic equation with $\mu = 0$ the order of convergence is $\gamma = 1$ for $\vartheta = 0$, 1, and $\gamma = 2$ for $\vartheta = 0.5$.

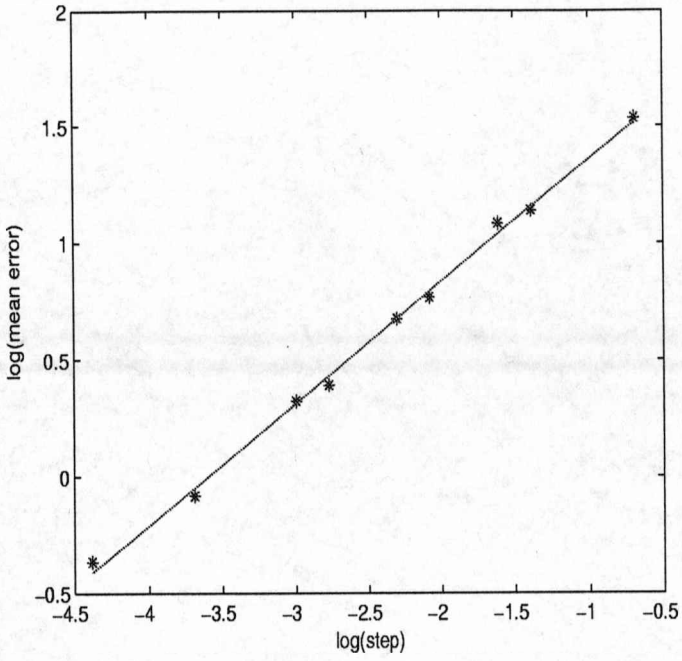


Figure 5.1: Variation of log(absolute mean endpoint error) with log(step length) for $\vartheta = 0$

5.3.2 Weak convergence

For weak convergence we are now looking at the absolute error between the expected value of the numerical solution and the expected true solution. We will again use the interval [0,1] and consider the error at the end point $t = 1$. For these experiments we will use the SDE

$$dY(t) = 2Y(t) + 0.1Y(t)dW(t), \qquad Y(0) = 1$$

as in [49] it is shown that $E(Y(1)) = e^2$. If we write

$$E_W = |E(Y_n) - E(Y(1))| = |E(Y_n) - e^2|$$

then the definition of weak convergence gives us

$$E_W \leq Ch^\gamma \tag{5.10}$$

or, by taking logarithms

$$\log(E_W) \leq \gamma \log(h) + \log(C). \tag{5.11}$$

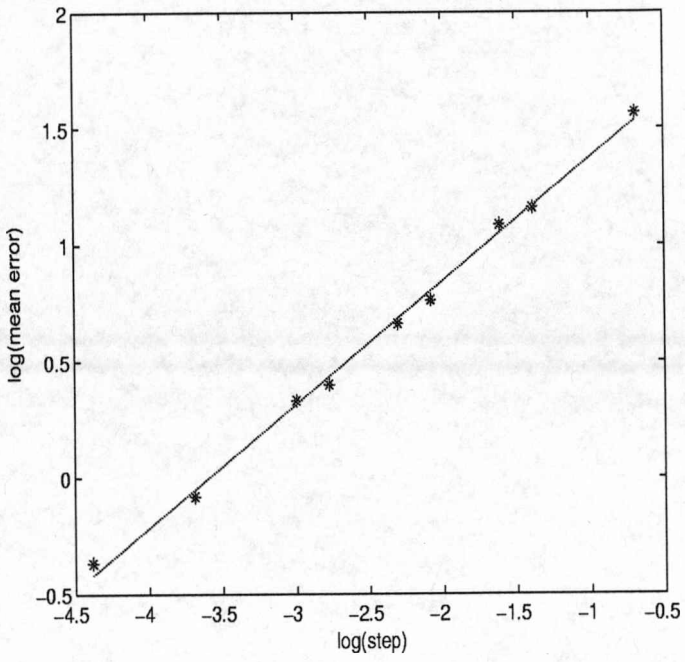


Figure 5.2: Variation of log(absolute mean endpoint error) with log(step length) for $\vartheta = 0.5$

Before actually investigating the weak convergence of the ϑ-methods, we confirm that, for this equation, they converge to $E(Y(1)) = e^2$ as $h \to 0$ for all values of ϑ.

Our equation is

$$dY(t) = 2Y(t) + 0.1Y(t)dW(t), \qquad Y(0) = 1$$

and the $\vartheta - methods$ give us the scheme

$$y_{n+1} = y_n + 2h[(1-\vartheta)y_n + \vartheta y_{n+1}] + 0.1\sqrt{(h)}\xi_n$$

where $\xi_n \in N(0,1)$, the standard normal distribution with zero mean and unit standard deviation. We can rearrange the scheme to get

$$y_{n+1} = \frac{1+2h(1-\vartheta)}{1-2h\vartheta}y_n + \frac{\sqrt{(h)}}{1-2h\vartheta}\xi_n$$

If we take expectactions, noting that $E[\xi_n] = 0$, we get

$$E[y_{n+1}] = \frac{1+2h(1-\vartheta)}{1-2h\vartheta}E[y_n]$$

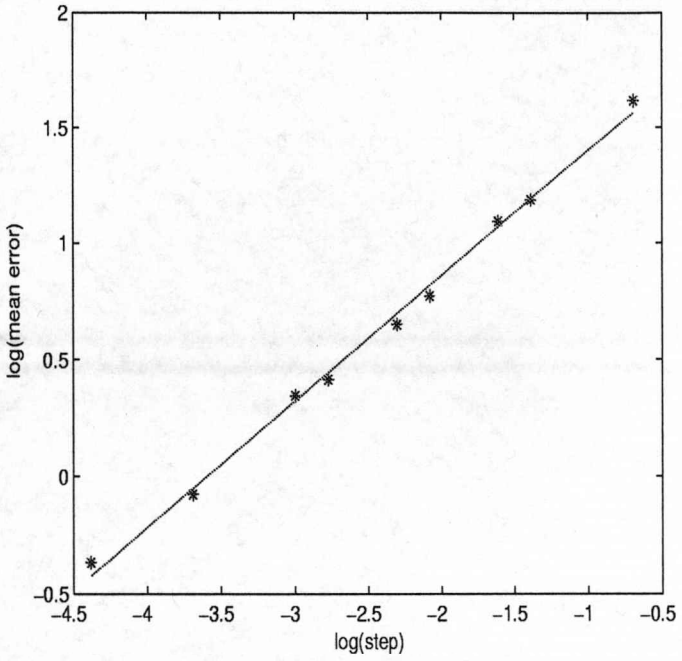


Figure 5.3: Variation of log(absolute mean endpoint error) with log(step length) for $\vartheta = 1$

Hence

$$E[y_n] = \left[\frac{1 + 2h(1-\vartheta)}{1 - 2h\vartheta}\right]^n E[y_0]$$

We have $y_0 = 1$, $y_n = y(1)$ and $nh = 1$ and so

$$\begin{aligned} E[y(1)] &= \left[\frac{1 + 2h(1-\vartheta)}{1 - 2h\vartheta}\right]^{1/h} \\ &\to \frac{e^{2(1-\vartheta)}}{e^{-2\vartheta}}, \qquad as\ h \to 0 \\ &= e^2, \end{aligned}$$

as predicted.

Once again we used MATLAB to simulate numerical solutions. For $\vartheta = 0$, 1 we used 20000 Brownian paths with increment $\delta t = \frac{1}{1600}$ to produce 20000 numerical solutions for each of the 8 step lengths $h = \frac{1}{800}, \frac{1}{400}, \frac{1}{200}, \frac{1}{160}, \frac{1}{100}, \frac{1}{50}, \frac{1}{25}, \frac{1}{20}$. We estimated $E(Y_n)$ for each step size by calculating the mean of the 20000 values of Y_n. We can now calculate E_W and plot the points for $\log(E_W)$ against $\log(h)$. Once again we found excellent linear fits and the plots, together with the least squares regression lines are shown in figures 5.4 and 5.5. For $\vartheta = 0$ the

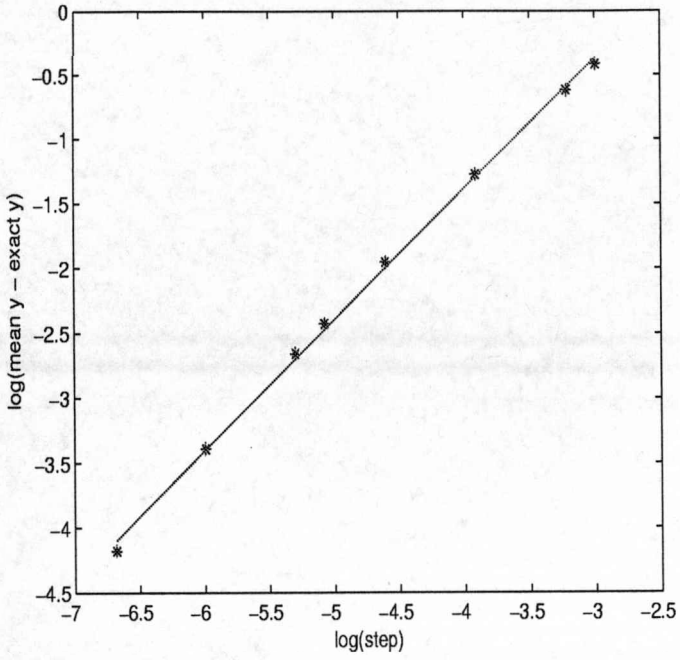


Figure 5.4: Variation of log(weak error) with log(step length) for $\vartheta = 0$

equation of the regression line is

$$\log(E_W) \approx 1.0120 \log(h) + 2.6637$$

which gives us

$$E_S \approx 14.3493 h^{1.0120}.$$

For $\vartheta = 1$ the equation of the regression line is

$$\log(E_W) \approx 0.9967 \log(h) + 2.7621$$

which gives us

$$E_S \approx 15.8331 h^{0.9967}.$$

We can see that for $\vartheta = 0, 1$ both methods have weak order of convergence $\gamma = 1$.

Initial experiments using the same scheme with $\vartheta = 0.5$ were inconclusive. However, when we increased the number of Brownian paths to 400000 and added the five additional step lengths $h = \frac{1}{1600}, \frac{1}{16}, \frac{1}{10}, \frac{1}{8}, \frac{1}{4}$ we produced the points shown in figure 5.6. It was noted that the six larger step lengths produce an excellent linear fit given by

$$\log(E_W) \approx 1.9460 \log(h) + 1.5428$$

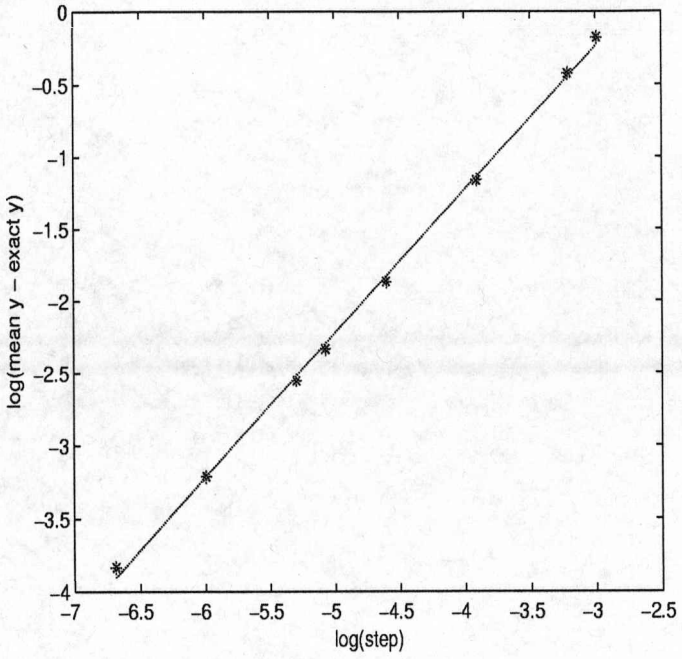


Figure 5.5: Variation of log(weak error) with log(step length) for $\vartheta = 1$

or

$$E_W \approx 4.6777 h^{1.9460},$$

showing weak order of convergence $\gamma \approx 2$. The regression line has been added to the figure. The weak error E_W for the smaller step lengths appear very small and fairly constant.

Further discussion on order of stochastic convergence can be found in [67, 51].

In conclusion, the convergence experiments in this section indicate that we are justified in using all three stochastic $\vartheta-$methods in our future work. For further analysis to the convergence of $\vartheta-$methods in particular we refer the reader to [50, 61, 9].

5.4 Methodology

In the following section we shall give details of our numerical results from running experiments. Our approach is as follows: we fix values for λ and μ and, for a fixed value of $\vartheta \in [0, 1]$, we choose a fixed step length $h = 1/N$ for some natural number N. Using Matlab (and its built-in random number generator), we simulate a large number of solution trajectories each for the same initial function and each solution is calculated over $[0, T]$ for some large value of T. For either the linear

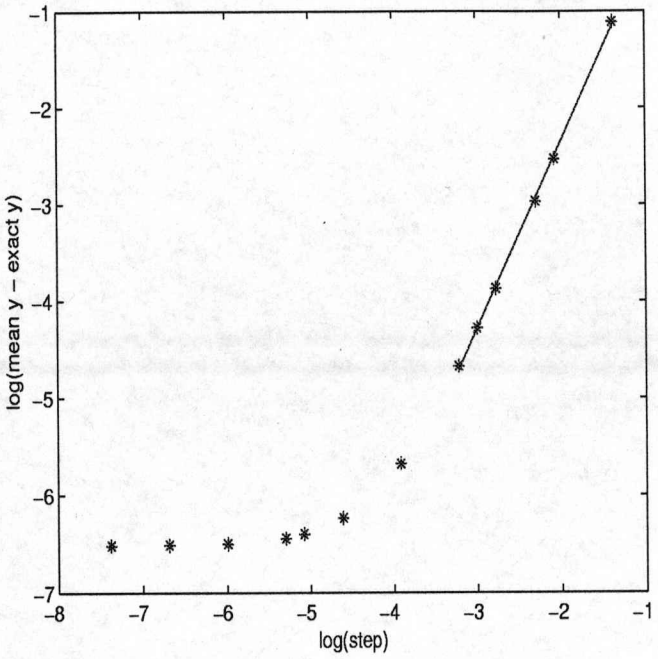


Figure 5.6: Variation of log(weak error) with log(step length) for $\vartheta = 0.5$

equation (4.3) or the nonlinear equation (4.4), for each solution trajectory Y we can calculate $S = \sup_{[T-\epsilon,T]}(|Y|)$ and we use this value to calculate $L = \frac{\log(S)}{T}$ which, for the linear equation at least, we regard as an estimate for the (local) Lyapunov exponent ([80]) based on the single solution trajectory Y. Thus, over a large number of simulations, we obtain, for each trajectory, a value of L. This gives us a large sample of estimates of the Lyapunov exponent and we consider the probability distribution of these values L. Of course, for the nonlinear equation (4.4) the limit given in (5.1) does not necessarily exist and we are not estimating a Lyapunov exponent, but instead the values of L that we calculate give us information about the dynamical behaviour of each solution trajectory.

We are interested in the following questions:

1. Can we identify the probability distribution of L values with some degree of confidence?

2. Can we detect how the values of $\vartheta, h, \lambda, \mu, \epsilon$ influence the distribution of L values?

3. Can we use the insights gained to give a clearer statement about stochastic bifurcations and to resolve some of the uncertainties identified in the P-bifurcation approach from [5].

5.5 Experimental results

Using the notation of the previous section, we fix $T = 5000$. We need to choose ϵ to enable us to estimate the lim sup in (5.1) so we know ϵ must be positive. Experimental work showed that $\epsilon = 50h$ gave reliable results. We are then able to present some results for various values of $\vartheta, h, \lambda, \mu$. In each case we report the least, the greatest, the mean and the standard deviation of the estimates of the Lyapunov exponent for the 500 solution paths we calculated. We produced data files for L using the following parameter values:

1: For the linear equation (4.3) we used 51 values of λ from -1.80 to -1.30 in steps of 0.01. (This range of values is roughly centred on $-\frac{\pi}{2}$). For each value of λ we took μ =0, 0.05, 0.1, 0.15, 0.2, 0.25 and 0.3. For each combination we applied the seven stepsizes h =0.5, 0.25, 0.2, 0.125, 0.1, 0.0625 and 0.05. Finally we used three values of ϑ = 0, 0.5 and 1. Consequently, we compiled 3x51x7x7 = 7497 values of L.

2: For the logistic equation (4.4) we restricted our experiments to μ =0.1 and h =0.1 only and took our range for λ as -1.85 to -1.45. We added extra linear experiments in order to make a comparison between the linear and logistic results at these parameter values. For the reasons stated in the following chapter we did not think it necessary to produce the same comprehensive set of results for this logistic equation.

Remark 5.5.1 We considered carefully our choices of T and ϵ and the number of solution paths. The natural changes one might make to ϵ and to the number of solution paths made little difference to our results and the values we used are a compromise between speed of calculation and the need for a large enough sample to provide consistent results. We also repeated our experiments for different finite values of T without significant changes in our conclusions. The choice of a value for T provides an interesting insight into the difference between numerical and analytical approaches. The analytical approach describes asymptotic behaviour of solutions as $T \to \infty$. However if a very large value of T is chosen then the values of the solution (even after all reasonable attempts at scaling) may be either zero or infinite (to machine precision). In other words, the limitations on the sizes of number (both large and small) that can be recorded accurately by the computer becomes a restriction on the ability to estimate accurately the value L for very large T. We believe that this is an important (and often neglected) factor in the use of simulations to show long term dynamical behaviour of solutions to equations. It was a factor observed previously in [79] who adopted a similar scaling formula for the calculation of approximate Lyapunov exponents that avoid these difficulties. However our motivation is to see how far numerical schemes reflect the underlying dynamics of the problem and therefore it is not our *primary purpose* simply to find a more efficient estimate for the values of Λ. To re-emphasise, the values of L describe the behaviour of solutions over

$[0, T]$ while the value of Λ provides information about the theoretical behaviour of solutions as $T \to \infty$. By considering the distribution of values of L one can deduce whether or not they can be used to give a reliable indication of the value of Λ.

We first need to consider the consistency of the sets of 500 values of L at each parameter setting. For the linear equation and the case $\vartheta = 0, h = 0.1, \mu = 0.1$ we have drawn histograms of the values of L for $\lambda = -1.65, -1.49, -1.34$. In all three cases we get bell-shaped curves. We have added normal distribution curves, with the same mean and standard deviation, to these histograms and we can see that there is a very close fit. These histograms are shown as figure 5.7. The Kolmogorov-Smirnov statistical test can be used to test if the distribution of L fits a normal distribution, [77]. The results of applying this test, using the SPSS statistical package, are given in table 5.1. A significance value of less than 0.05 would give evidence that the distributions are not normal. All three of our histograms give p-values greater than 0.200 and hence we conclude that there is no evidence to say that our distributions of L are not normal. Similar results were obtained for a random selection of other parameter settings, but not all 7497 cases were considered.

	Kolmogorov-Smirnov[a] normality test		
λ	Statistic	degrees of freedom	Significance
-1.65	0.020	500	0.200*
-1.49	0.025	500	0.200*
-1.34	0.019	500	0.200*

Table 5.1: SPSS table for normal distribution fit of the histograms
* This is a lower bound of the true significance,
[a] Lilliefors significance correction.

Looking at the range of values of L for the three histograms we should not be too surprised at these values. At $\lambda = -1.65$ the phenomenological experiments indicated that the solutions to the linear equation becomes unbounded. All 500 results of L at this value of λ are positive. At $\lambda = -1.34$ the phenomenological experiments indicated that the solutions to the linear equation converge to zero. All 500 results of L at this value of λ are negative. Finally, the phenomenological experiments put $\lambda = -1.49$ into a region of uncertainty and the values of L span zero, indicating that some solutions become unbounded and some converge, depending upon the Brownian path used. The mean value of L is -0.001328, very close to zero.

The three tables 5.2, 5.3 and 5.4 give examples of our data collected for the linear equation (4.3) for $\vartheta = 0, 0.5$ and 1 respectively. The tables all use a stepsize of $h = 0.1$ and noise coefficient $\mu = 0.1$. The values in the table are given to

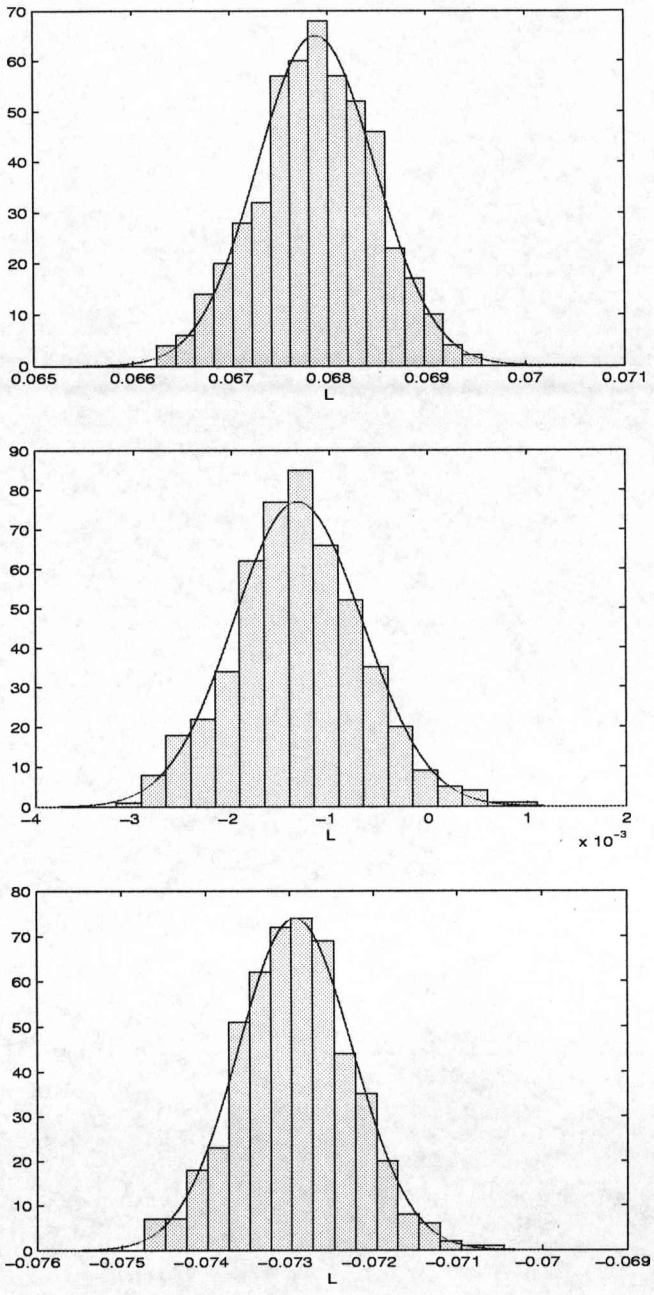


Figure 5.7: Histogram of the 500 values of L for $\vartheta = 0$, $\mu = 0.1$ and stepsize $h = 0.1$
Top:$\lambda = -1.65$
Middle:$\lambda = -1.49$
Bottom: $\lambda = -1.34$

6 decimal places, but we note that the data files were produced in MATLAB's long format mode to 15 decimal places.

All three tables 5.2, 5.3 and 5.4 show fairly similar results. The values of L_{min}, L_{mean} and L_{max} all appear to decrease at a similar steady rate for all three values of ϑ. There is also very little variation in the standard deviation of L across the values of λ and ϑ, ranging approximately from -0.0007 at $\lambda = -1.3$ to -0.0006 at $\lambda = -1.8$. However, this is not a constant increase as λ increases. Figure 5.8 clearly shows the variation of the standard deviation of L, confirming very little variation.

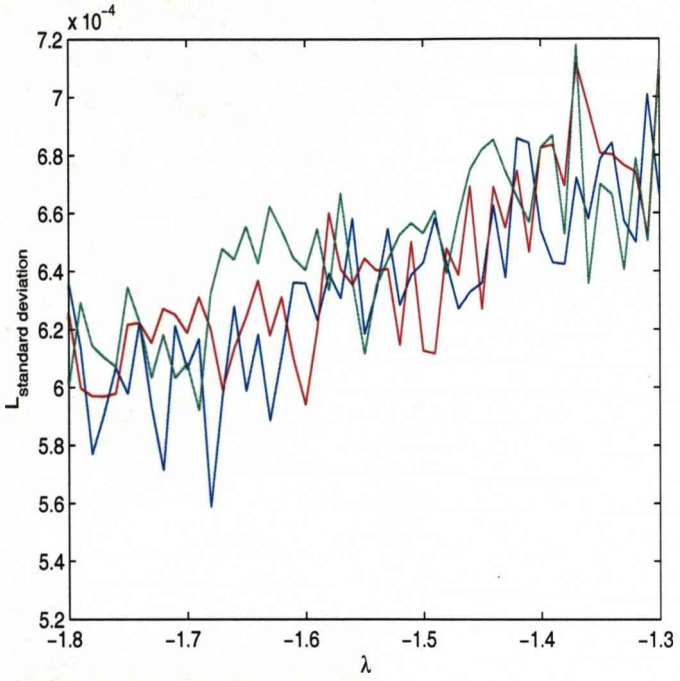


Figure 5.8: Variation of the standard deviation of L for $h = 0.1$, $\mu = 0.1$
$\vartheta = 0$: Blue $\vartheta = 0.5$: Red $\vartheta = 1$: Green

We now consider the effect on L of varying the step length. We use the value $\lambda = -1.57$ and keep μ at 0.1. Tables 5.5, 5.6 and 5.7 to give the results for the Euler Maruyama, Semi-implicit trapezium and Euler rules respectively.

We note at this point that tables 5.5, 5.6 and 5.7 show no particular pattern in the values of the stardard deviations as h varies.

Finally in this chapter we consider the effect on L of varying the noise coefficient. We again use the value $\lambda = -1.57$ and keep h at 0.1. Tables 5.8, 5.9 and 5.10 give the results for the Euler Maruyama, Semi-implicit trapezium and Euler rules respectively.

λ	L_{min}	L_{mean}	L_{max}	$L_{Standard\ deviation}$
-1.80	0.125696	0.127203	0.129467	0.000636
-1.79	0.121622	0.123360	0.125211	0.000614
-1.78	0.117992	0.119548	0.121777	0.000577
-1.77	0.113750	0.115720	0.117549	0.000591
-1.76	0.110093	0.111871	0.113710	0.000607
-1.75	0.106262	0.107970	0.109446	0.000598
-1.74	0.102413	0.104070	0.105727	0.000623
-1.73	0.097770	0.100131	0.102324	0.000593
-1.72	0.094544	0.096156	0.097865	0.000572
-1.71	0.090348	0.092147	0.094711	0.000621
-1.70	0.086109	0.088214	0.090290	0.000607
-1.69	0.082100	0.084158	0.085716	0.000617
-1.68	0.078722	0.080134	0.081711	0.000559
-1.67	0.074561	0.076060	0.077643	0.000596
-1.66	0.068938	0.071973	0.073805	0.000628
-1.65	0.066243	0.067877	0.069558	0.000599
-1.64	0.061837	0.063713	0.065542	0.000618
-1.63	0.057768	0.059592	0.061584	0.000589
-1.62	0.053654	0.055401	0.057312	0.000611
-1.61	0.049075	0.051211	0.052994	0.000636
-1.60	0.045156	0.046968	0.048924	0.000636
-1.59	0.040832	0.042669	0.044500	0.000623
-1.58	0.036722	0.038417	0.040472	0.000639
-1.57	0.032551	0.034093	0.036211	0.000631
-1.56	0.027892	0.029741	0.032026	0.000658
-1.55	0.023601	0.025392	0.027206	0.000619
-1.54	0.018852	0.021045	0.022701	0.000633
-1.53	0.014776	0.016588	0.018400	0.000655
-1.52	0.010178	0.012110	0.014359	0.000628
-1.51	0.005969	0.007695	0.009414	0.000639
-1.50	0.001550	0.003210	0.005319	0.000643
-1.49	-0.003188	-0.001328	0.001100	0.000658
-1.48	-0.007740	-0.005871	-0.003872	0.000643
-1.47	-0.012154	-0.010492	-0.008343	0.000627
-1.46	-0.016888	-0.015106	-0.013391	0.000633
-1.45	-0.022140	-0.019750	-0.017955	0.000636
-1.44	-0.026084	-0.024429	-0.022442	0.000663
-1.43	-0.031114	-0.029131	-0.027093	0.000638
-1.42	-0.035916	-0.033894	-0.031982	0.000686
-1.41	-0.040636	-0.038619	-0.036827	0.000684
-1.40	-0.045154	-0.043385	-0.041458	0.000654
-1.39	-0.050181	-0.048274	-0.046355	0.000643
-1.38	-0.054813	-0.053148	-0.051360	0.000642
-1.37	-0.060171	-0.058072	-0.055989	0.000672
-1.36	-0.065142	-0.062943	-0.060652	0.000658
-1.35	-0.070033	-0.067943	-0.066053	0.000679
-1.34	-0.074742	-0.072934	-0.070458	0.000684
-1.33	-0.080270	-0.078028	-0.075960	0.000657
-1.32	-0.085634	-0.083084	-0.081031	0.000650
-1.31	-0.090096	-0.088245	-0.085680	0.000701
-1.30	-0.095153	-0.093354	-0.090911	0.000666

Table 5.2: L values for $h = 0.1, \mu = 0.1, \vartheta = 0$

λ	L_{min}	L_{mean}	L_{max}	$L_{Standard\ deviation}$
-1.80	0.094212	0.096078	0.098069	0.000626
-1.79	0.090461	0.092120	0.093989	0.000600
-1.78	0.086519	0.088170	0.090245	0.000597
-1.77	0.082166	0.084174	0.086281	0.000597
-1.76	0.078209	0.080156	0.082316	0.000598
-1.75	0.074231	0.076072	0.077923	0.000622
-1.74	0.070114	0.071966	0.073960	0.000622
-1.73	0.065739	0.067858	0.070242	0.000616
-1.72	0.062022	0.063689	0.066115	0.000627
-1.71	0.057747	0.059541	0.061628	0.000625
-1.70	0.053523	0.055394	0.057368	0.000618
-1.69	0.049490	0.051209	0.053097	0.000631
-1.68	0.045191	0.047022	0.048744	0.000620
-1.67	0.041138	0.042756	0.044575	0.000599
-1.66	0.036702	0.038469	0.040719	0.000614
-1.65	0.032254	0.034219	0.035982	0.000624
-1.64	0.028143	0.029867	0.031920	0.000637
-1.63	0.023922	0.025516	0.027158	0.000618
-1.62	0.019345	0.021132	0.023198	0.000631
-1.61	0.014952	0.016751	0.018489	0.000610
-1.60	0.010644	0.012383	0.014320	0.000594
-1.59	0.005962	0.007908	0.009949	0.000620
-1.58	0.001157	0.003460	0.005791	0.000660
-1.57	-0.003083	-0.001092	0.000744	0.000641
-1.56	-0.007360	-0.005620	-0.003231	0.000635
-1.55	-0.012035	-0.010137	-0.008254	0.000644
-1.54	-0.016631	-0.014719	-0.012811	0.000640
-1.53	-0.021334	-0.019322	-0.017068	0.000640
-1.52	-0.026138	-0.023987	-0.022032	0.000614
-1.51	-0.030700	-0.028672	-0.026152	0.000650
-1.50	-0.034971	-0.033385	-0.031606	0.000612
-1.49	-0.039650	-0.038079	-0.036292	0.000612
-1.48	-0.044789	-0.042862	-0.041242	0.000648
-1.47	-0.049391	-0.047600	-0.045760	0.000639
-1.46	-0.055027	-0.052430	-0.050497	0.000669
-1.45	-0.059569	-0.057287	-0.055127	0.000627
-1.44	-0.064067	-0.062176	-0.060003	0.000669
-1.43	-0.069160	-0.067024	-0.064703	0.000655
-1.42	-0.074139	-0.072020	-0.070262	0.000674
-1.41	-0.078877	-0.076987	-0.074741	0.000647
-1.40	-0.083900	-0.082015	-0.079894	0.000682
-1.39	-0.088938	-0.087059	-0.084840	0.000684
-1.38	-0.094348	-0.092098	-0.090305	0.000669
-1.37	-0.099259	-0.097252	-0.095216	0.000712
-1.36	-0.104802	-0.102433	-0.100083	0.000697
-1.35	-0.109357	-0.107604	-0.105410	0.000681
-1.34	-0.114909	-0.112802	-0.110494	0.000680
-1.33	-0.119989	-0.118048	-0.115331	0.000677
-1.32	-0.125559	-0.123368	-0.121536	0.000674
-1.31	-0.130544	-0.128733	-0.126832	0.000653
-1.30	-0.136436	-0.134108	-0.132036	0.000713

Table 5.3: L values for $h = 0.1, \mu = 0.1, \vartheta = 0.5$

λ	L_{min}	L_{mean}	L_{max}	$L_{Standard\ deviation}$
-1.80	0.063275	0.065329	0.067415	0.000601
-1.79	0.059346	0.061138	0.063047	0.000629
-1.78	0.055302	0.056933	0.058531	0.000614
-1.77	0.050758	0.052655	0.055021	0.000610
-1.76	0.046678	0.048380	0.050480	0.000607
-1.75	0.042383	0.044137	0.046021	0.000635
-1.74	0.038156	0.039802	0.041698	0.000623
-1.73	0.033310	0.035497	0.037811	0.000603
-1.72	0.029403	0.031161	0.033175	0.000618
-1.71	0.024575	0.026833	0.028436	0.000603
-1.70	0.020426	0.022424	0.024566	0.000608
-1.69	0.016204	0.017980	0.019871	0.000592
-1.68	0.011701	0.013548	0.016302	0.000632
-1.67	0.007071	0.009076	0.010990	0.000648
-1.66	0.002679	0.004584	0.006735	0.000644
-1.65	-0.002031	0.000028	0.001969	0.000655
-1.64	-0.006274	-0.004455	-0.002459	0.000643
-1.63	-0.010920	-0.009046	-0.007175	0.000662
-1.62	-0.015825	-0.013675	-0.011893	0.000654
-1.61	-0.020576	-0.018294	-0.016273	0.000645
-1.60	-0.024579	-0.022968	-0.021100	0.000640
-1.59	-0.029464	-0.027656	-0.025100	0.000655
-1.58	-0.034453	-0.032372	-0.030495	0.000633
-1.57	-0.039270	-0.037144	-0.034987	0.000667
-1.56	-0.044396	-0.041870	-0.039304	0.000636
-1.55	-0.048557	-0.046663	-0.044615	0.000612
-1.54	-0.053737	-0.051468	-0.049301	0.000635
-1.53	-0.058034	-0.056328	-0.053986	0.000644
-1.52	-0.063040	-0.061211	-0.058626	0.000653
-1.51	-0.067981	-0.066201	-0.063764	0.000657
-1.50	-0.073253	-0.071171	-0.069049	0.000653
-1.49	-0.078185	-0.076134	-0.074247	0.000661
-1.48	-0.082835	-0.081073	-0.079490	0.000639
-1.47	-0.088188	-0.086150	-0.083954	0.000659
-1.46	-0.093395	-0.091190	-0.089378	0.000675
-1.45	-0.098941	-0.096357	-0.094304	0.000682
-1.44	-0.103377	-0.101528	-0.099790	0.000685
-1.43	-0.108852	-0.106717	-0.104828	0.000675
-1.42	-0.114175	-0.111885	-0.109738	0.000665
-1.41	-0.119479	-0.117120	-0.115102	0.000657
-1.40	-0.124474	-0.122438	-0.120428	0.000683
-1.39	-0.129823	-0.127792	-0.125205	0.000687
-1.38	-0.134978	-0.133090	-0.130805	0.000653
-1.37	-0.140284	-0.138473	-0.136256	0.000718
-1.36	-0.145886	-0.143916	-0.141710	0.000636
-1.35	-0.151404	-0.149448	-0.146932	0.000670
-1.34	-0.156885	-0.154941	-0.153179	0.000666
-1.33	-0.162351	-0.160524	-0.158463	0.000640
-1.32	-0.168096	-0.166081	-0.163768	0.000679
-1.31	-0.173361	-0.171697	-0.169692	0.000650
-1.30	-0.179505	-0.177375	-0.174561	0.000711

Table 5.4: L values for $h = 0.1, \mu = 0.1, \vartheta = 1$

h	L_{min}	L_{mean}	L_{max}	$L_{Standard\ deviation}$
0.5	0.138783	0.140232	0.138027	0.000524
0.25	0.076646	0.078879	0.080973	0.000583
0.2	0.062613	0.064688	0.066410	0.000599
0.125	0.040349	0.042078	0.044135	0.000618
0.1	0.032551	0.034093	0.036211	0.000631
0.0625	0.020187	0.021782	0.024462	0.000660
0.05	0.015351	0.017638	0.019735	0.000657

Table 5.5: L values for $\lambda = -1.57, \mu = 0.1, \vartheta = 0$

h	L_{min}	L_{mean}	L_{max}	$L_{Standard\ deviation}$
0.5	-0.036212	-0.034328	-0.032719	0.000569
0.25	-0.010057	-0.008390	-0.006546	0.000630
0.2	-0.007139	-0.005221	-0.002734	0.000671
0.125	-0.004045	-0.001839	0.000095	0.000649
0.1	-0.003083	-0.001092	0.000744	0.000641
0.0625	-0.001971	-0.000242	0.002589	0.000689
0.05	-0.001944	-0.000064	0.001631	0.000643

Table 5.6: L values for $\lambda = -1.57, \mu = 0.1, \vartheta = 0.5$

h	L_{min}	L_{mean}	L_{max}	$L_{Standard\ deviation}$
0.5	-0.240643	-0.238684	-0.236666	0.000699
0.25	-0.103207	-0.101423	-0.099657	0.000636
0.2	-0.080953	-0.078734	-0.076817	0.000641
0.125	-0.048771	-0.047080	-0.044789	0.000652
0.1	-0.039270	-0.037144	-0.034987	0.000669
0.0625	-0.024748	-0.022632	-0.020262	0.000676
0.05	-0.020067	-0.017942	-0.015648	0.000659

Table 5.7: L values for $\lambda = -1.57, \mu = 0.1, \vartheta = 1$

μ	L_{min}	L_{mean}	L_{max}	$L_{Standard\ deviation}$
0.00	-0.037800	-0.037800	-0.037800	0.000000
0.05	-0.038820	-0.037779	-0.036694	0.000320
0.10	-0.039270	-0.037144	-0.034987	0.000669
0.15	-0.038895	-0.036017	-0.033175	0.000961
0.20	-0.038327	-0.034502	-0.030356	0.001333
0.25	-0.037770	-0.032522	-0.027188	0.001599
0.30	-0.035337	-0.030171	-0.024711	0.001905

Table 5.8: L values for $\lambda = -1.57, h = 0.1, \vartheta = 1$

μ	L_{min}	L_{mean}	L_{max}	$L_{Standard\ deviation}$
0.00	-0.001892	-0.001892	-0.001892	0.000000
0.05	-0.002551	-0.001694	-0.000551	0.000306
0.10	-0.003083	-0.001092	0.000744	0.000641
0.15	-0.002730	-0.000074	0.002629	0.000976
0.20	-0.002713	0.001308	0.005354	0.001272
0.25	-0.001366	0.003008	0.007558	0.001498
0.30	0.000093	0.005142	0.010712	0.001899

Table 5.9: L values for $\lambda = -1.57, h = 0.1, \vartheta = 1$

μ	L_{min}	L_{mean}	L_{max}	$L_{Standard\ deviation}$
0.00	0.033362	0.033362	0.033362	0.000000
0.05	0.032736	0.033545	0.034341	0.000305
0.10	0.032551	0.034093	0.036211	0.000631
0.15	0.032097	0.035027	0.037929	0.000983
0.20	0.032977	0.033383	0.033755	0.000126
0.25	0.034509	0.038014	0.043767	0.001512
0.30	0.033907	0.039933	0.046412	0.001838

Table 5.10: L values for $\lambda = -1.57, h = 0.1, \vartheta = 1$

The tables 5.8, 5.9 and 5.10 clearly indicate a linear relationship between the standard deviation of the 500 values of L and the noise coefficient μ. Figure 5.9 highlights this linear relationship. It appears that doubling the noise coefficient will double the spread of values for the 500 results of L, which is not surprising.

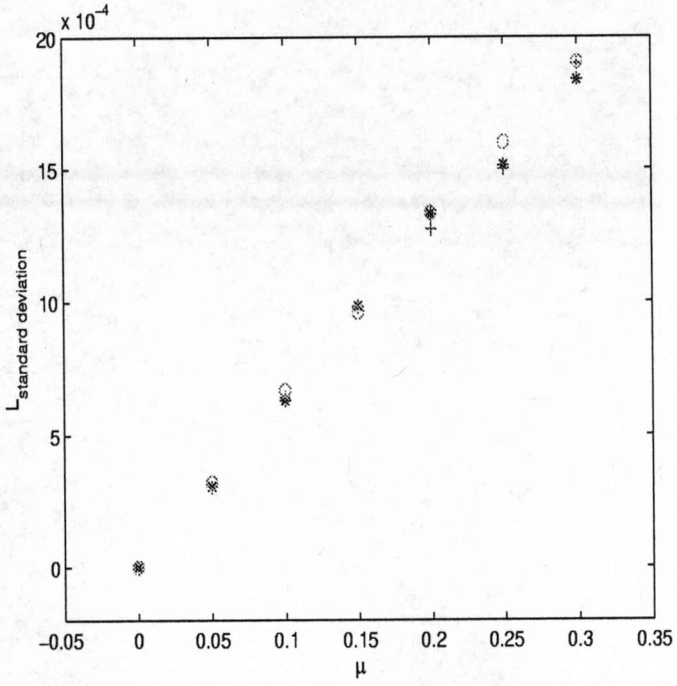


Figure 5.9: Variation of $L_{standard\ deviation}$ at different noise coefficients for $h = 0.1$, $\lambda = -1.57$
$\vartheta = 0$: * $\vartheta = 0.5$: + $\vartheta = 1$: o

5.6 Review of conclusions

Our motivation in this initial investigation has been to consider (using a more precise approach than that adopted in the paper [5]) how reliably numerical schemes perform when they are used to predict changes in the dynamical behaviour of solutions to stochastic delay differential equations.

We can draw the following conclusions, with reference to the questions in the chapter's introduction.

1. The different numerical schemes we have tried for varying values of ϑ give a consistent story but the detailed results are different for each choice of scheme, ϑ, step length, h, and noise coefficient μ.

2. Different (finite) values of T lead to essentially the same results. Theoretical results about the asymptotic behaviour as $t \to \infty$ do not *necessarily* give accurate information about what will be seen in a solution over a *finite* time interval.

At the beginning of the chapter we stated that our principal concern is to find out whether numerical methods can be relied upon to give accurate and precise information about changes in the qualitative behaviour of the solutions to equations. In our previous work we showed that this was indeed the case for deterministic equations undergoing a bifurcation. However this chapter has shown clearly that insights gained from numerical simulations of stochastic equations should be treated with much greater caution. In particular, one *must not* expect to draw reliable conclusions about the long-term dynamical behaviour of solutions to a stochastic delay differential equation using results from simulations based on a single numerical scheme and step length over a finite interval.

Chapter 6

The search for a formula: I

[4] The objective of this chapter is to start the development of a formula to predict the parameter relationships at the bifurcation of the numerical solution to our test equation. This proves to be quite a complex task and it will occupy this and the following two chapters as we refine our approach. Here we aim to find a formula, using all three ϑ-methods, for L, our estimate of the lyapunov exponent, in terms of the parameter λ for a particular step size. We will then solve the equation $L = 0$ to compute a D-bifurcation value for λ. We can then compare this value with the P-bifurcation value we produced in chapter 4.

6.1 A possible linear relationship

We continue to take $T = 5000$ which is sufficiently large to give consistent results without being so large that the experiments take excessive time and take ϵ so that for our experiments we took the final 50 values of Y_n for calculating S. This value of ϵ also gave reliable results. The graphs of the minimum, mean and maximum values of L for the linear and logistic equations, using $h = 0.1$ and $\mu = 0.1$, are shown for $\vartheta = 0, 0.5, 1$ in figures 6.1, 6.2.

It can be seen that the graphs for the linear equations produce curves which are extremely close to straight lines. The logistic equations give the same approximate straight lines for $\lambda > \lambda_{bif}$. For values of $\lambda < \lambda_{bif}$ the graphs of L against λ for the logistic equations become unpredictable. This change in the distribution of L for the two equations is not surprising. A condition for the linearisation of the logistic equation is $Y(t) \approx 0$, and this condition no longer holds beyond the bifurcation point (where $|Y(t)|$ increases). For both the linear and logistic equation we find that $|Y(t)| \to 0$ for $\lambda > \lambda_{bif}$ so we would expect the values of L to be very close for this region. However, for $\lambda < \lambda_{bif}$ $|Y(t)| \to \infty$ for the linear

[4]The content and the approach in the next two chapters appeared as 'Predicting changes in dynamical behaviour in solutions to stochastic delay differential equations' in Communications on Pure and Applied Analysis, Volume 5, Number 2, June 2006, pp. 367-382, [71].

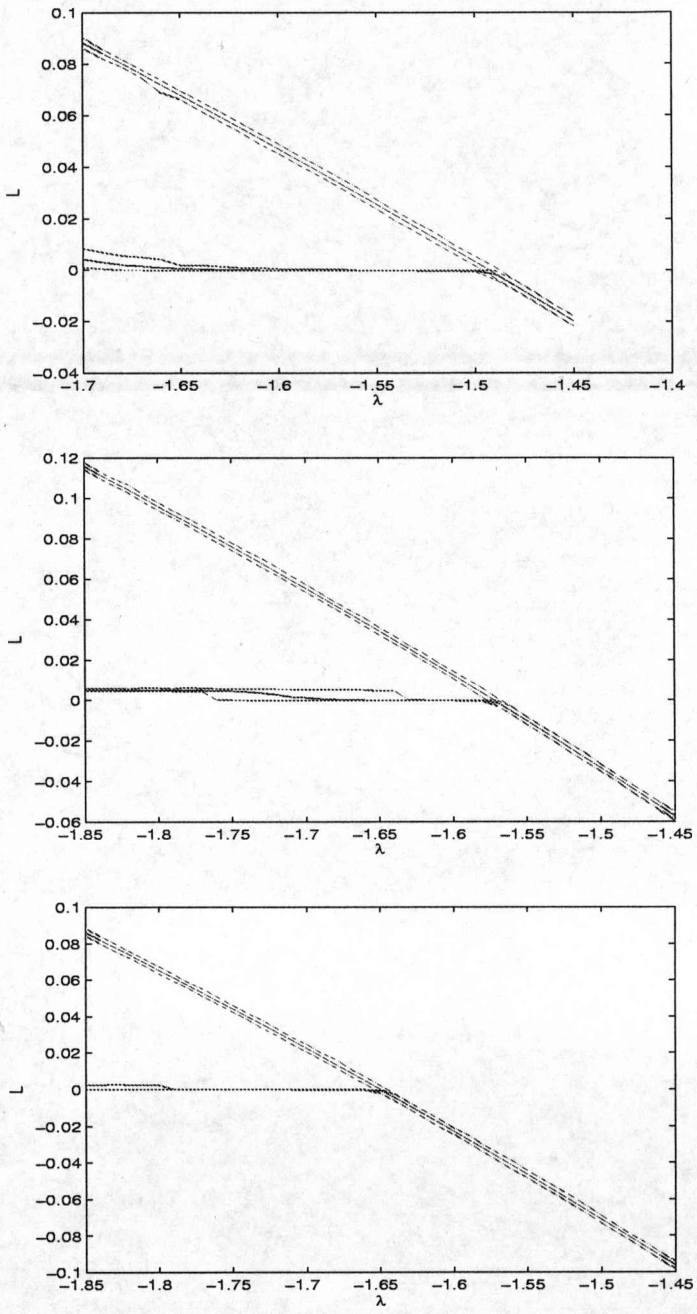


Figure 6.1: Minimum, mean and maximum values of $L, h = 0.1, \mu = 0.1$
Top: $\vartheta = 0$
Middle: $\vartheta = 0.5$
Bottom: $\vartheta = 1$

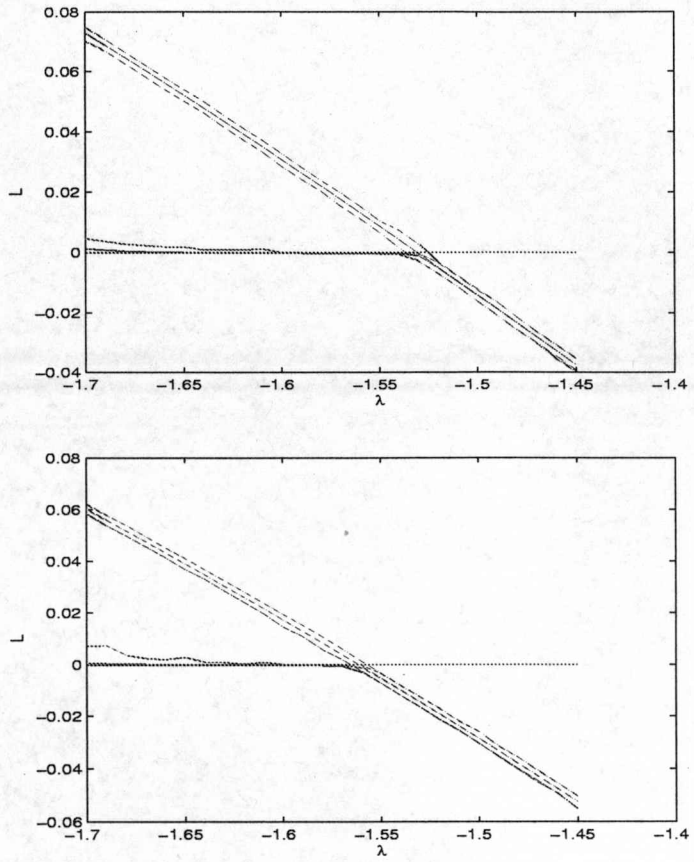


Figure 6.2: Minimum, mean and maximum values of $L, \mu = 0.1$
Top: $\vartheta = 0, h = 0.05$
Bottom: $\vartheta = 0, h = 0.01$

equation, but oscillates for the logistic equation. Different distributions of L can now be expected. As this is an initial search for a formula we will restrict most of the remaining work in this chapter to a stepsize $h = 0.1$ and noise coefficient $\mu = 0.1$. However, we will conclude the chapter with a collection of graphs that demonstrate the generalised results with a mix of stepsize and noise coefficients.

Following the comments on the linearisation of the logistic equation and the clear indication in figures 6.1, 6.2 of the lack of uniformity of the positive L values we will restrict all the further work in this thesis to the linear equation.

We converted the MATLAB data files into SPSS[1] files for our statistical analysis. Using SPSS we can calculate the least squares regression lines for the

[1]Statistical Package for Social Scientists, a commercial statistics computer package for analysing data, marketed by SPSS inc.

ϑ		Regression line	R^2
0	Min	$L = -0.431\lambda - 0.646$	0.999
	Mean	$L = -0.431\lambda - 0.644$	1
	Max	$L = -0.430\lambda - 0.640$	0.999
0.5	Min	$L = -0.432\lambda - 0.683$	0.999
	Mean	$L = -0.432\lambda - 0.680$	0.999
	Max	$L = -0.432\lambda - 0.678$	0.999
1	Min	$L = -0.456\lambda - 0.755$	0.999
	Mean	$L = -0.455\lambda - 0.752$	0.999
	Max	$L = -0.454\lambda - 0.750$	0.999

Table 6.1: Equations of regression lines of L against λ

ϑ	D-bifurcation			P-bifurcation
	Min	Mean	Max	
0	-1.4988	-1.4942	-1.4884	-1.4946
0.5	-1.5810	-1.5741	-1.5694	-1.5740
1	-1.6557	-1.6527	-1.6520	-1.6516

Table 6.2: Linear regression values of λ_{bif}

linear equations in the form $L = a\lambda + b$ for the minimum, mean and maximum lines for each value of ϑ. SPSS will give us the values of a and b, and the correlation coefficient R. We can also calculate the coefficient of determination, R^2, a measure of the reliability of the straight line fits, [77]. R^2 is known to vary from 0 to 1, with 1 giving a perfect fit. The equations are given in table 6.1. All nine equations have an R^2 value of at least 0.999, which indicate an excellent fit. An example of the lines is shown in figure 6.3.

We are detecting the parameter value at which the Lyapunov exponent changes sign. We can use our least squares equations to estimate the values of λ at which L changes sign and compare these values with the values of λ_{bif} found earlier for our P-bifurcations. Table 6.2 shows these values.

Table 6.2 shows that the D-bifurcation estimates correspond very closely to the P-bifurcations found in section 4.2 for the deterministic equation. However, we would not expect equal values for the P- and the D-bifurcations. The differences observed reflect the fact that λ_{bif} will depend upon the magnitude of the coefficient of the multiplicative noise as well as on the choice of numerical scheme and step length. We can see from the figures that, for smaller step lengths, λ_{bif} becomes closer to $-\frac{\pi}{2}$ as in the deterministic case. We have $\mu = 0$ for the P-bifurcations in section 4.2 and $\mu = 0.1$ for the D-bifurcations here.

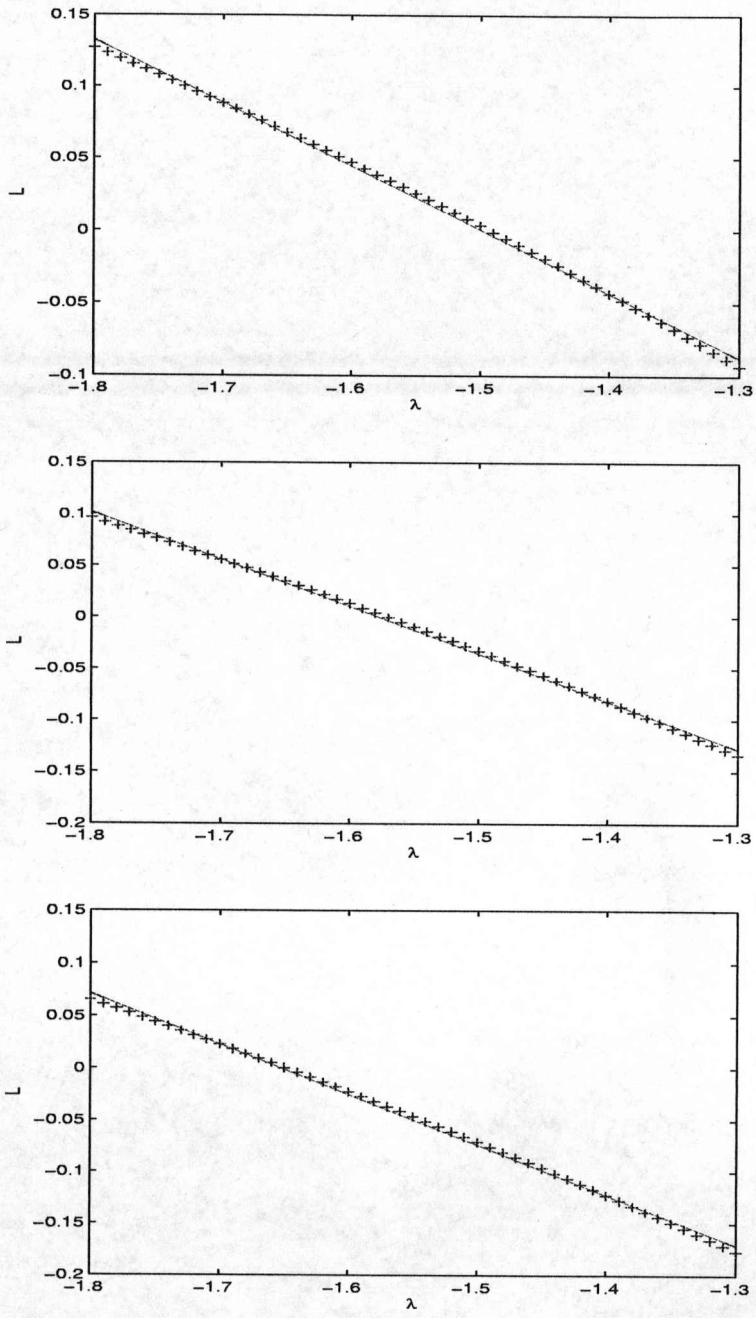


Figure 6.3: Mean L, with the linear regression line, for $\mu = 0.1$
Top: $\vartheta = 0$
Middle: $\vartheta = 0.5$
Bottom: $\vartheta = 1$

6.2 Further graphs

We finish this chapter with a collection of further graphs, figure 6.4, showing the near linear relationship between Ł and λ for some different parameter values. These graphs show consistent results across a range of parameter values.

6.3 Conclusions

In this chapter we have clearly seen that the value of L depends upon λ in a predictable way, and that a linear relationship gives us a good initial approximation. We have also shown that the solution of the equation $L(\lambda) = 0$ will provide us with a good estimate of the bifurcation value of λ.

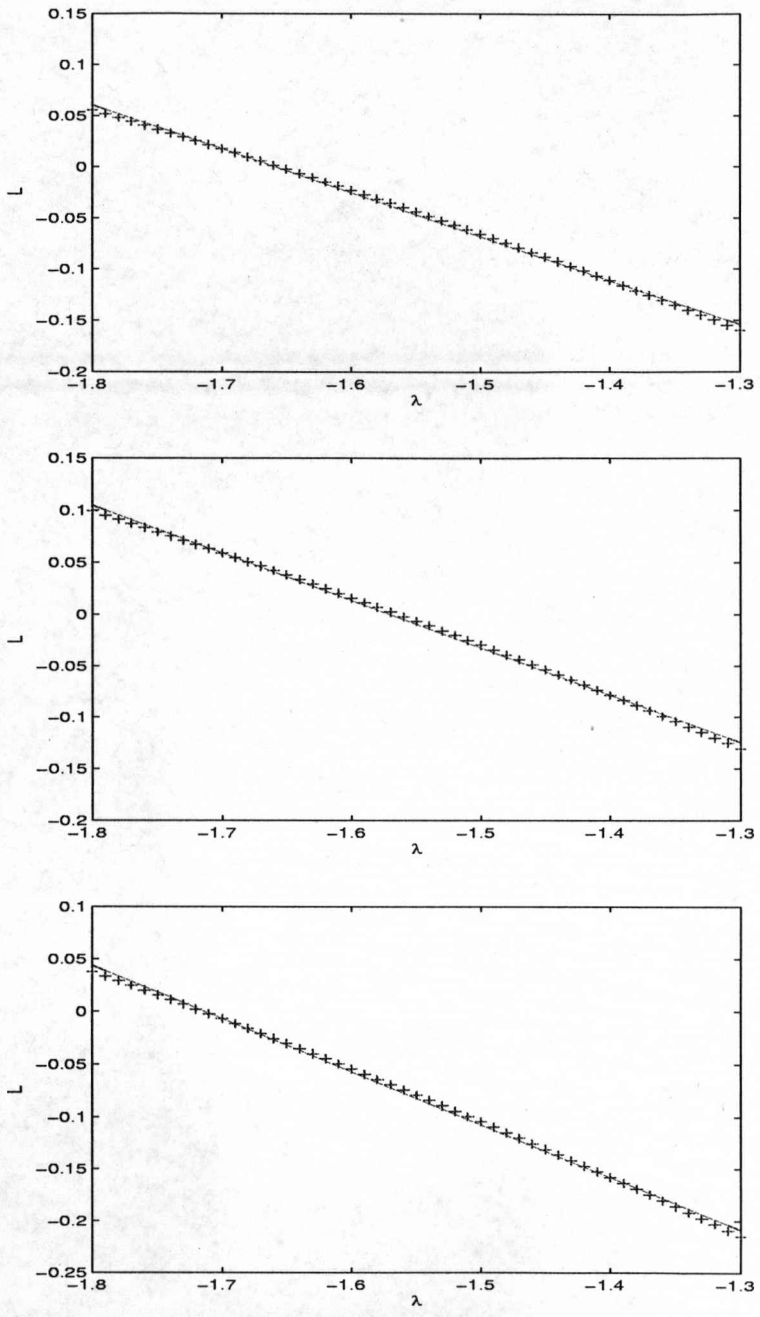


Figure 6.4: Mean L against λ with the equations of the regression lines
Top: $\vartheta = 0.5, \mu = 0.0, h = 0.5$ Line $L = -0.428227\lambda - 0.710665$
Middle: $\vartheta = 0.5, \mu = 0.2, h = 0.05$ Line $L = -0.457951\lambda - 0.719847$
Bottom: $\vartheta = 1, \mu = 0.3, h = 0.2$ Line $L = -0.503450\lambda - 0.862837$

Chapter 7

The search for a formula: II

[5] The graphs in the previous chapter showed us that there is a relationship between L_{mean} and λ. Our first attempts to find a formula for this relationship, using a linear fit, gave us reasonable results. In this chapter we aim to improve upon this initial approach.

7.1 Quadratic relationships

7.1.1 Variation of L with λ

Fixing h

A visual inspection of the regression lines indicates that the data follow a graph that is very slightly concave upwards. We will continue to use the statistical package SPSS to analyse our data. Using regression analysis we found that the mean values of L appear to follow a near perfect quadratic curve. We used SPSS to find the least squares quadratic formulae for the cases $\mu = 0$, 0.1, 0.2, 0.3 with $h = 0.1$ for all three of our ϑ values. These formulae are given in table 7.1, together with their coefficients of determination, R^2. All twelve coefficients $R^2 = 1$, confirming our perfect fits. We can show these fits with L_{mean} plotted as points, with the regression curves superimposed, as figures 7.1, 7.2.

Figure 7.1 shows a series of graphs for a fixed ϑ value and varying μ. We can see that changes in μ move the curves in very small increments to the right. In fact, the curves are so close that they are almost identical. We can also show the curves for fixed μ and varying ϑ. These curves are shown in figure 7.2 and we can see that these curves are further apart. By completing and comparing both sets of figures we can see that all twelve curves are close to being parallel.

[5] Much of the content and the approach in this chapter appeared as 'Using approximations to Lyapunov exponents to predict changes in dynamical behaviour in numerical solutions to stochastic delay differential equations', Algorithms for Approximation, [35].

ϑ	μ	Quadratic equation	R^2
0	0.0	$L = -0.13619\lambda^2 - 0.86265\lambda - 0.98525$	1
	0.1	$L = -0.13558\lambda^2 - 0.86002\lambda - 0.98189$	1
	0.2	$L = -0.13415\lambda^2 - 0.85355\lambda - 0.97303$	1
	0.3	$L = -0.13183\lambda^2 - 0.84286\lambda - 0.95835$	1
0.5	0.0	$L = -0.14166\lambda^2 - 0.89887\lambda - 1.06387$	1
	0.1	$L = -0.14111\lambda^2 - 0.89647\lambda - 1.06065$	1
	0.2	$L = -0.13883\lambda^2 - 0.88725\lambda - 1.04941$	1
	0.3	$L = -0.13498\lambda^2 - 0.87179\lambda - 1.03078$	1
1	0.0	$L = -0.14998\lambda^2 - 0.94961\lambda - 1.15914$	1
	0.1	$L = -0.14974\lambda^2 - 0.94805\lambda - 1.15638$	1
	0.2	$L = -0.14903\lambda^2 - 0.94349\lambda - 1.14832$	1
	0.3	$L = -0.14630\lambda^2 - 0.93113\lambda - 1.13139$	1

Table 7.1: Quadratic regression curves, with $h = 0.1$.

With our dynamical approach for finding D-bifurcations we need to find the parameter values at which $L = 0$. For our first tabulated case, $\vartheta = 0$, $h = 0.1$, $\mu = 0$, we need to solve the equation

$L = -0.13619\lambda^2 - 0.86265\lambda - 0.98525 = 0$

to find the λ values at bifurcation. This gives us two roots, -4.8392 and -1.4950, and from observations of the shape of the curves we clearly need the greater root, -1.4950. Repeating for the other eleven equations we can summarise our results in table 7.2. We have added the corresponding P-bifurcation values for the deterministic equation to this table. We would expect the P-bifurcation values to be close to the D-bifurcation values at $\mu = 0$. Inspection of the table confirms this fact.

Fixing μ

If we now keep $\vartheta = 0$ and set $\mu = 0.1$ we can find the variation in L with different values of stepsize $h = 0.5, 0.25, 0.2, 0.125, 0.1, 0.0625, 0.05$. Using SPSS we can find the seven quadratic regression curves, and these are given in table 7.3. The coefficients of determination $R^2 = 1$, to three decimal places, in all seven cases, showing perfect fits.

The values of L are plotted for each h in figure 7.3, with the regression curves superimposed. The curves are clearly close to parallel.

Solving the equations $L = 0$, and taking the relevant root, we can find the value of λ at the D-bifurcations.The values, together with the P-bifurcation values for the deterministic equation ($\mu = 0$) found by the techniques in chapter 4, are shown in table 7.4.

Clearly we would not expect the same values for the P-bifurcation and the

ϑ	Quadratic μ	D-bifurcation value of λ	P-bifurcation value of λ
0	0.0	-1.4950	-1.4946
	0.1	-1.4932	
	0.2	-1.4879	
	0.3	-1.4727	
0.5	0.0	-1.5740	-1.5740
	0.1	-1.5722	
	0.2	-1.5670	
	0.3	-1.5584	
1	0.0	-1.6513	-1.6516
	0.1	-1.6495	
	0.2	-1.6441	
	0.3	-1.6352	

Table 7.2: Quadratic regression values of λ_{bif}, with $h = 0.1$.

h	Quadratic equation	R^2
0.5	$L = -0.118576\lambda^2 - 0.745596\lambda - 0.738039$	1
0.25	$L = -0.127790\lambda^2 - 0.809591\lambda - 0.877140$	1
0.2	$L = -0.129967\lambda^2 - 0.824641\lambda - 0.909579$	1
0.125	$L = -0.134155\lambda^2 - 0.851038\lambda - 0.963252$	1
0.1	$L = -0.135577\lambda^2 - 0.860018\lambda - 0.981888$	1
0.0625	$L = -0.138171\lambda^2 - 0.875473\lambda - 1.012050$	1
0.05	$L = -0.138904\lambda^2 - 0.880319\lambda - 1.022054$	1

Table 7.3: L_{mean} against λ with $\mu = 0.1$, $\vartheta = 0$.

h	D-bifurcation value of λ	P-bifurcation value of λ
0.5	-1.2308	-1.2361
0.25	-1.3872	-1.3892
0.2	-1.4214	-1.4232
0.125	-1.4748	-1.4763
0.1	-1.4932	-1.4946
0.0625	-1.5212	-1.5226
0.05	-1.5307	-1.5321

Table 7.4: Quadratic regression values of λ_{bif}, $\vartheta = 0$, $\mu = 0.1$.

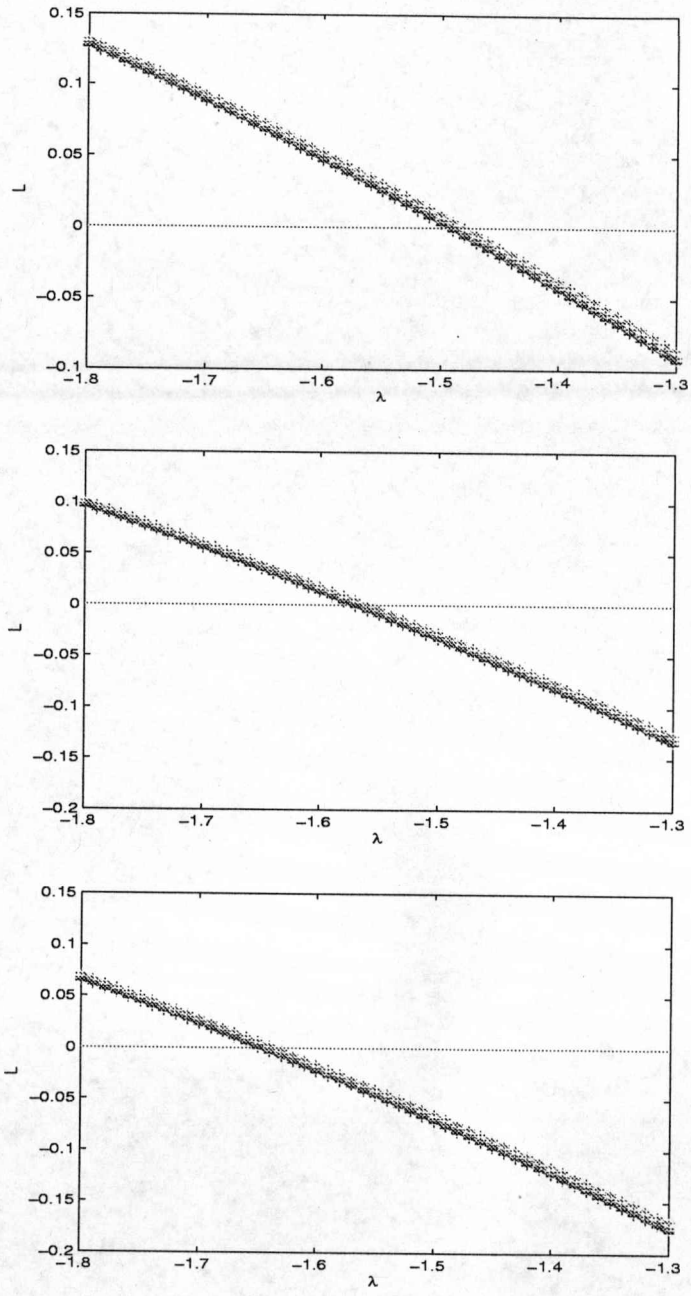


Figure 7.1: Mean L, with the quadratic regression curves for $\mu = 0, 0.1, 0.2, 0.3$ left to right on each graph, with $h = 0.1$
Top: $\vartheta = 0$
Middle: $\vartheta = 0.5$
Bottom: $\vartheta = 1$

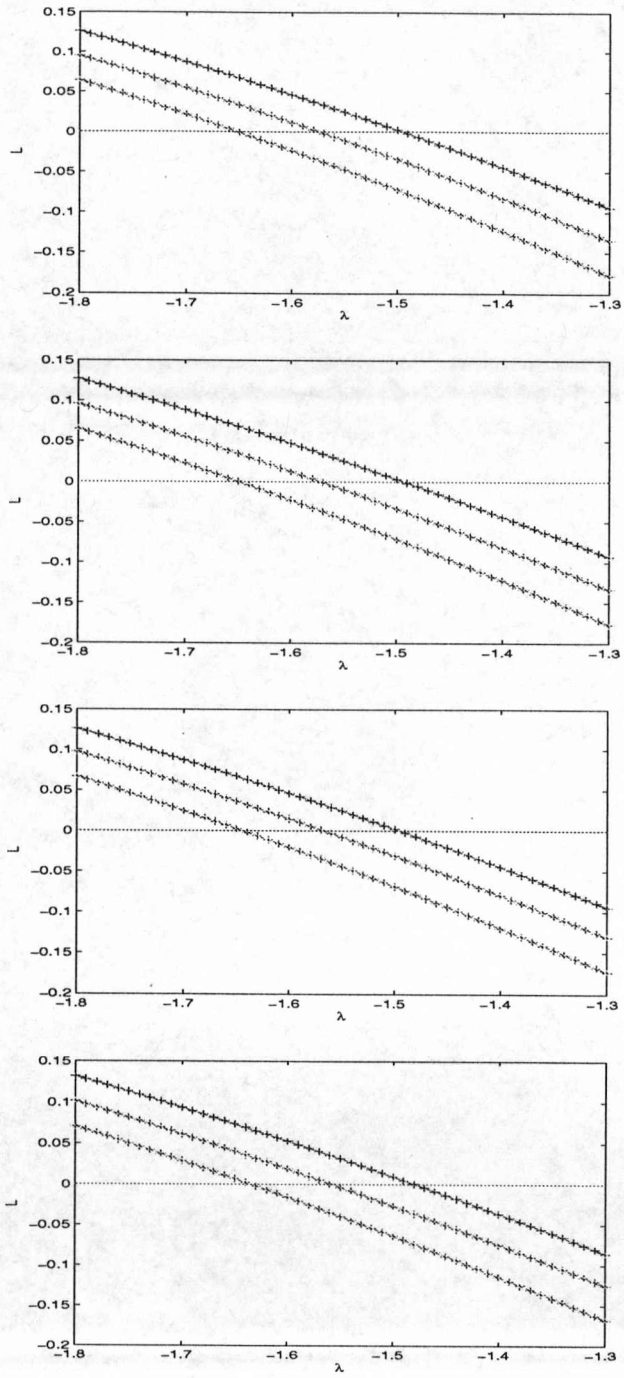


Figure 7.2: Mean L, with the quadratic regression curves for left to right $\vartheta = 1, 0.5, 0$ on each graph, and from top to bottom $\mu = 0, 0.1, 0.2, 0.3$ with $h = 0.1$.

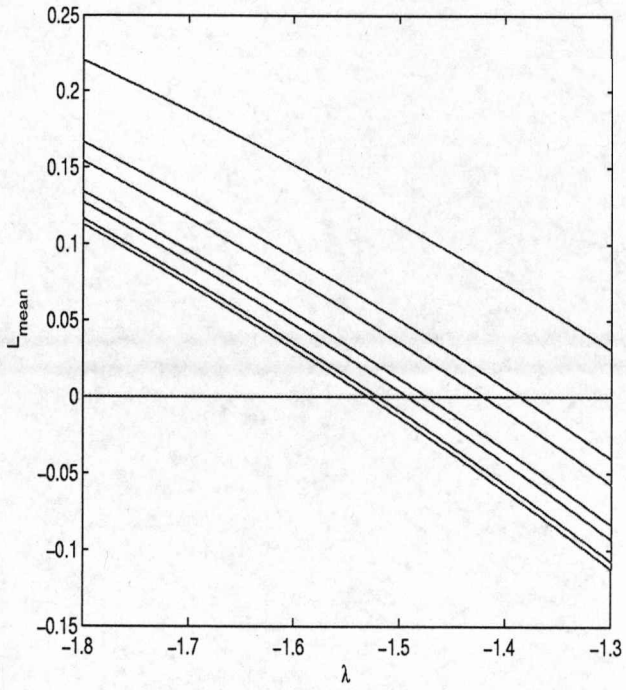


Figure 7.3: Mean L, with the quadratic regression curves for $\vartheta = 0$, $\mu = 0.1$. Top to bottom $h = 0.5, 0.25, 0.2, 0.125, 0.1, 0.0625, 0.05$

D-bifurcation (we have different values for μ), but we would expect close values. The table confirms this.

7.1.2 Variation of L with h

Finally we can investigate how L_{mean} varies with h for fixed λ.

We already know that for a fixed μ and h we obtain an excellent fit for L_{mean} as a quadratic function of λ. We can now investigate the fit for L_{mean} if we fix λ and μ. We need to begin by determining an appropriate model to choose. We base this on the following insight: For the deterministic case we know that it can be shown (theoretically) that the numerical bifurcation point approximates the exact bifurcation to the order of the method. Therefore it makes sense to base our models on the order of the numerical methods in use. It has been shown in [49] that the Euler Maruyama method has strong order of convergence $\gamma = 0.5$ and weak order of convergence $\gamma = 1$, and for $\vartheta = 0.5$ we have a weak order $\gamma = 2$. Consequently we looked for a relationship using $h^{\frac{1}{2}}$, h and h^2 as the dominant terms. The results of experiments with different combinations in the models are given in table 7.5 for the case $\lambda = -1.50$, together with the

Equation	R
$L_{mean} = 0.4795h^2 - 0.00074$	.954
$L_{mean} = 0.2641\sqrt{h} - 0.07847$	.995
$L_{mean} = 0.2835h - 0.02506$	.998
$L_{mean} = 0.170250h + 0.107215\sqrt{h} - 0.047085$	.99993
$L_{mean} = -0.143577h^2 + 0.362442h - 0.031696$	.999997
$L_{mean} = -0.1215886h^2 + 0.3323459h + 0.0170465\sqrt{h} - 0.0341816$	1
$L_{mean} = -0.2140995h^{1.5} + 0.4660743h - 0.0180839\sqrt{h} - 0.0309053$	1

Table 7.5: Simple regression formulae for $\vartheta = 0$ and $\lambda = -1.50$

λ	Quadratic equation	R^2
-1.8	$L = -0.119031h^2 + 0.303866h + 0.097956$	1
-1.7	$L = -0.126717h^2 + 0.322257h + 0.057182$	1
-1.6	$L = -0.133858h^2 + 0.341172h + 0.014109$	1
-1.5	$L = -0.143577h^2 + 0.362442h - 0.031696$	1
-1.4	$L = -0.152311h^2 + 0.383909h - 0.080388$	1
-1.3	$L = -0.164308h^2 + 0.408497h - 0.132665$	1

Table 7.6: L_{mean} against h with $\mu = 0.1$, $\vartheta = 0$.

correlation coefficients, R. The closer the value of R is to 1, the better the fit.

The conclusions here need to be interpreted with care and there is scope for further experimentation to reach a completely firm conclusion. One must bear in mind the fact that, by introducing additional complexity in the model, one may obtain falsely accurate results. Both of the final two equations provide an almost perfect fit of the data points ($R = 1$ to 10 significant figures) but there is a much stronger dependency on the terms in h and of higher order than on the term in $\sqrt{h}$. Both this observation, and further experimentation with other values of λ has led us to conclude that we should use the quadratic model in our analysis but this decision is provisional and needs to be reviewed when further analytical and/or numerical evidence becomes available. We can now examine this relationship further. Taking $\vartheta = 0$, $\mu = 0.1$ we can plot the graph of L_{mean} against h for a range of values of λ, using our seven values of $h = 0.5, 0.25, 0.2, 0.125, 0.1, 0.0625, 0.05$. Again, using SPSS, we find that we get perfect quadratic fits and the quadratic regression curves are given in table 7.6.

Figure 7.4 shows our points together with these regression curves superimposed. Once again all of the curves visually appear parallel.

For completeness we can solve all seven quadratic equations $L = 0$, to find the values of h that give the bifurcation at the six values of λ used. These values are shown in table 7.7. For $\lambda = -1.8, -1.7, -1.6$ the roots are negative and, as

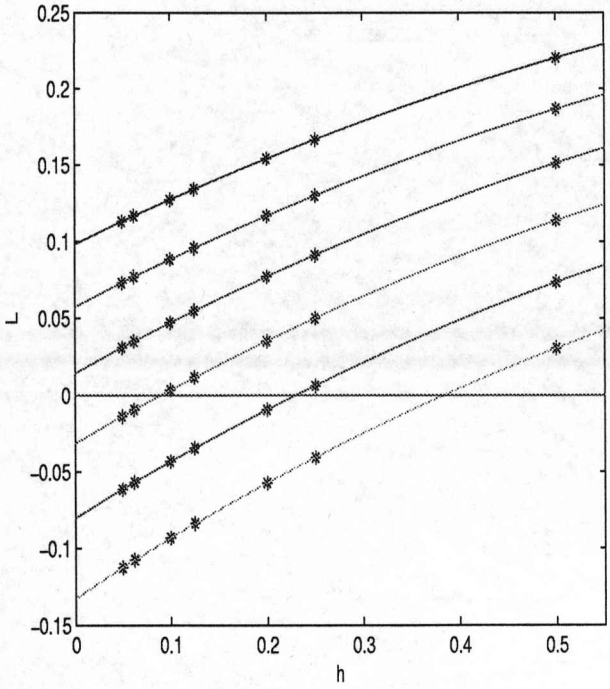


Figure 7.4: Mean L, with the quadratic regression curves for $\vartheta = 0$, $\mu = 0.1$. Top to bottom λ=-1.8, -1.7, -1.6, -1.5, -1.4, -1.3

$h > 0$, they are not feasible. This indicates that these λ give solutions to our test equation that all tend to infinity for any Brownian path.

7.2 Conclusions

In this chapter we have shown that our values for L and, in particular L_{mean} give a consistent picture of the bifurcation points of the solutions to our equation, with excellent sets of regression curves as we vary the equation and numerical method parameters. For each ϑ we can find values of L_{mean} for a combination of values for the equation parameters λ and μ and the numerical method stepsize h. We have shown there are a number of excellent quadratic regression curve fits for L_{mean} in terms of either λ or h when the other parameter values are fixed. In particular, we have produced quadratic relationships for

1. L_{mean} in terms of λ for fixed h. A change in the ϑ-method gives a set of parallel curves, while further sets of parallel curves are produced for changes in the value of μ.

λ	D-bifurcation value of h
-1.8	No value
-1.7	No value
-1.6	No value
-1.5	0.0907
-1.4	0.2305
-1.3	0.3841

Table 7.7: Quadratic regression values of h, $\mu = 0.1$, $\vartheta = 0$.

2. L_{mean} in terms of h for fixed λ. Different values of λ give us a set of parallel curves. Again, a change in the ϑ-method gives a further set of parallel curves. We have also shown that further sets of parallel curves are produced for changes in the value of μ, but, as h is usually taken as a simple fraction of the equation delay (τ=1), this analysis was thought to be less useful.

In each case we have solved the quadratic equation $L_{mean} = 0$ to find the appropriate parmeter value at the D-bifurcation. Comparisons, where possible, have been made with the P-bifurcation values from chapter 4, producing expected similarities in results. This all appears to confirm the suitability of our approach to analysing the bifurcations in the solutions to our test equation and this analysis is taken forward to a higher level in the following chapter.

Chapter 8

Further formulae and their implications

8.1 Introduction

[6] We now aim to pull together the separate analyses of the preceding chapters.

8.2 Multivariate quadratic relationships

In the previous chapter we found quadratic regression fits for L_{mean} in terms of λ and h, with μ fixed, separately. We will now investigate the relationship with λ and h together. If we plot the values of L_{mean} against λ and h for $\vartheta = 0, \mu = 0.1$ we see, in figure 8.1, a smooth surface. As in the previous chapters we can use SPSS to find a second order regression formula for L_{mean} in terms of λ and h. Tables 8.1, 8.2 and 8.3 list the formulae for the 21 cases we considered. In every case the correlation coefficient $R \approx 1$, indicating near perfect fits[7].

We can add the surface given by the regression formula for $\vartheta = 0, \mu = 0.1$ to the plot of points and the combined graph is shown in figure 8.2.

In the two dimensional case we found that all of our regression curves were parallel for different values of μ. We can now investigate this situation in three dimensions. Figure 8.3 shows the surfaces for $\vartheta = 0$ with $\mu = 0$ and $\mu = 0.3$ and the two surfaces clearly look parallel. Figure 8.4 shows the effect of changing the ϑ-method. It shows the surfaces for $\vartheta = 0$, $\vartheta = 0.5$ and $\vartheta = 1$ with $\mu = 0.1$. In

[6] The content and the approach in this chapter has been accepted for publication in the proceedings to the "Second international workshop on analysis and numerical approximations of singular problems" as "Noise-induced changes to the behaviour of semi-implicit Euler methods for stochastic delay equations undergoing bifurcation", [36].

[7] In order to simplify the algebra we omitted the term in λh. We considered the improvement in R from 0.9999886 to 0.9999892 to be insignificant. However, this approach could be investigated at a later stage.

μ	Equation	R
0.00	$L_{mean} = -.1324985\lambda^2 - .1382667h^2 - .8378776\lambda + .3516786h - .9896782$	.999
0.05	$L_{mean} = -.1323627\lambda^2 - .1386083h^2 - .8372673\lambda + .3520501h - .9888930$	.999
0.10	$L_{mean} = -.1318815\lambda^2 - .1394995h^2 - .8352395\lambda + .3531157h - .9864279$	.999
0.15	$L_{mean} = -.1311747\lambda^2 - .1410140h^2 - .8321481\lambda + .3549305h - .9825467$	.999
0.20	$L_{mean} = -.1301694\lambda^2 - .1433929h^2 - .8277879\lambda + .3576682h - .9771360$	.999
0.25	$L_{mean} = -.1288605\lambda^2 - .1465803h^2 - .8221417\lambda + .3613273h - .9701922$	.999
0.30	$L_{mean} = -.1274131\lambda^2 - .1506332h^2 - .8157371\lambda + .3660038h - .9621623$	.999

Table 8.1: Regression formulae for $\vartheta = 0$

μ	Equation	R
0.00	$L_{mean} = -.1385855\lambda^2 - .1337073h^2 - .8835123\lambda - .0019131h - 1.0460225$	1.000
0.05	$L_{mean} = -.1384194\lambda^2 - .1340008h^2 - .8827959\lambda - .0013900h - 1.0451492$	1.000
0.10	$L_{mean} = -.1378855\lambda^2 - .1347155h^2 - .8805295\lambda + .0001020h - 1.0424380$	1.000
0.15	$L_{mean} = -.1371017\lambda^2 - .1360919h^2 - .8770965\lambda + .0027371h - 1.0382274$	1.000
0.20	$L_{mean} = -.1358635\lambda^2 - .1382031h^2 - .8718543\lambda + .0065807h - 1.0320209$	1.000
0.25	$L_{mean} = -.1342600\lambda^2 - .1410728h^2 - .8651043\lambda + .0116786h - 1.0240963$	1.000
0.30	$L_{mean} = -.1323592\lambda^2 - .1450999h^2 - .8570765\lambda + .0183135h - 1.0146818$	1.000

Table 8.2: Regression formulae for $\vartheta = 0.5$

μ	Equation	R
0.00	$L_{mean} = -.1613891\lambda^2 - .3028416h^2 - 1.0160353\lambda - .3373565h - 1.1974603$	.997
0.05	$L_{mean} = -.1610697\lambda^2 - .3023180h^2 - 1.0147701\lambda - .3367636h - 1.1961002$	.997
0.10	$L_{mean} = -.1603796\lambda^2 - .3009093h^2 - 1.0118034\lambda - .3348000h - 1.1926925$	.997
0.15	$L_{mean} = -.1593837\lambda^2 - .2984645h^2 - 1.0073553\lambda - .3315786h - 1.1874168$	.997
0.20	$L_{mean} = -.1577823\lambda^2 - .2951257h^2 - 1.0004679\lambda - .3269734h - 1.1795373$	.997
0.25	$L_{mean} = -.1555875\lambda^2 - .2909900h^2 - 0.9911997\lambda - .3209129h - 1.1691346$	.997
0.30	$L_{mean} = -.1528499\lambda^2 - .2864202h^2 - 0.9797878\lambda - .3131905h - 1.1564873$	.997

Table 8.3: Regression formulae for $\vartheta = 1$

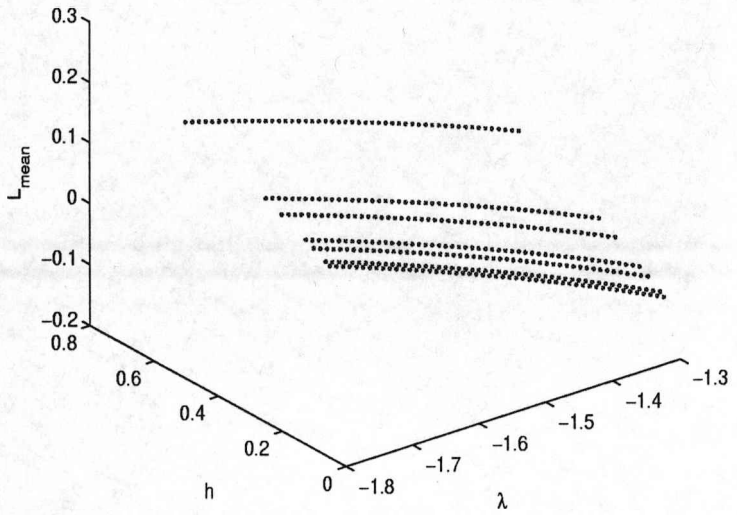


Figure 8.1: L_{mean} against λ and h for $\mu = 0.1$

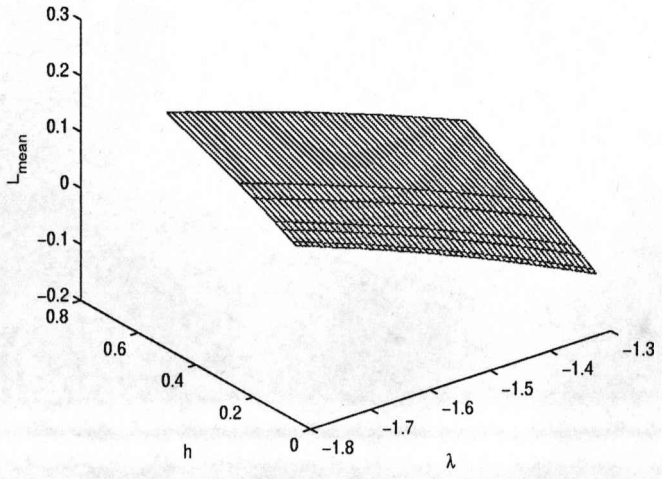


Figure 8.2: L_{mean} against λ and h for $\mu = 0.1$

this case the cross section along constant h are parallel, but those along constant λ are not.

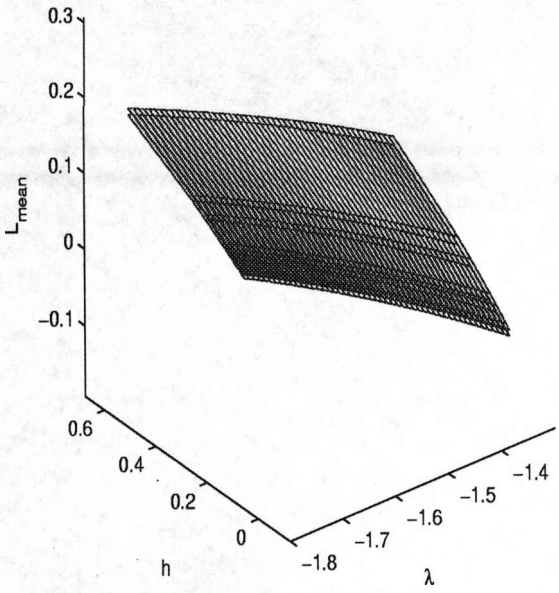


Figure 8.3: L_{mean} against λ and h for $\vartheta = 0$, with $\mu = 0$ and $\mu = 0.3$

We are interested in the parameter values of λ at $L_{mean} = 0$, as this will give us a value for λ_{bif} at the bifurcation point. Adding the plane $L_{mean} = 0$ to our initial figure 8.1 gives a clear picture of the situation with figure 8.5,

8.3 Results

We can write each equation for $L_{mean} = 0$ as

$$a\lambda^2 + b\lambda + c + dh + eh^2 = 0 \tag{8.1}$$

We can solve the equation for λ in terms of increasing powers of h.
First, for convenience, we let $D^2 = b^2 - 4ac$.
Using the quadratic formula , we obtain

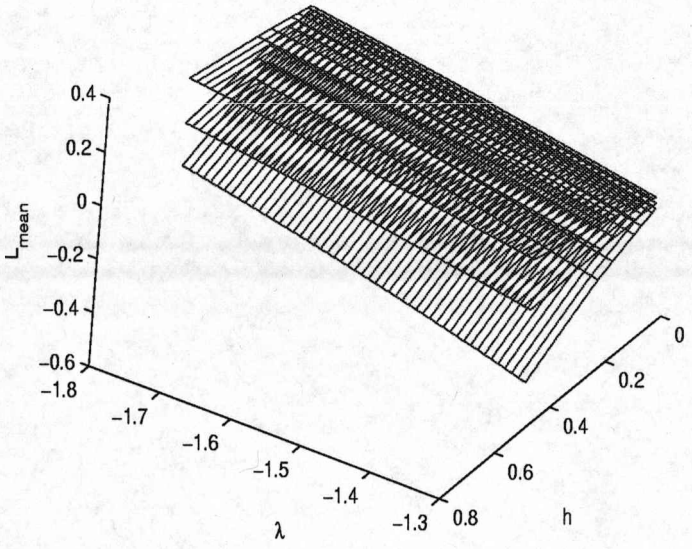


Figure 8.4: L_{mean} against λ and h for $\vartheta = 0, 0.5, 1$ with $\mu = 0.1$

$$\begin{aligned} \lambda &= \frac{-b \pm \sqrt{b^2 - 4a(c + dh + eh^2)}}{2a} \\ &= \frac{-b \pm D\sqrt{1 - 4a(dh + eh^2)/D^2}}{2a} \end{aligned} \tag{8.2}$$

Now, if we have

$$-D^2 \leq 4a(dh + eh^2) \leq D^2 \tag{8.3}$$

we can expand equation (8.2) in terms of h. With reference to Figure 8.2, we take the larger root of equation (8.2) which becomes

$$\begin{aligned} \lambda &= \frac{-b - D[1 - \frac{1}{2}4a(dh + eh^2)/D^2 - \frac{1}{8}16a^2(dh + eh^2)^2/D^4 + \ldots}{2a} \\ &= \frac{(-b - D)}{2a} + \frac{(dh + eh^2)}{D} + \frac{a(dh + eh^2)}{D^3} + \ldots \\ &= \frac{(-b - D)}{2a} + \frac{d}{D}h + (\frac{e}{D} + \frac{ad^2}{D^3})h^2 + \ldots \textit{terms in } h^3 \textit{ and higher} \end{aligned} \tag{8.4}$$

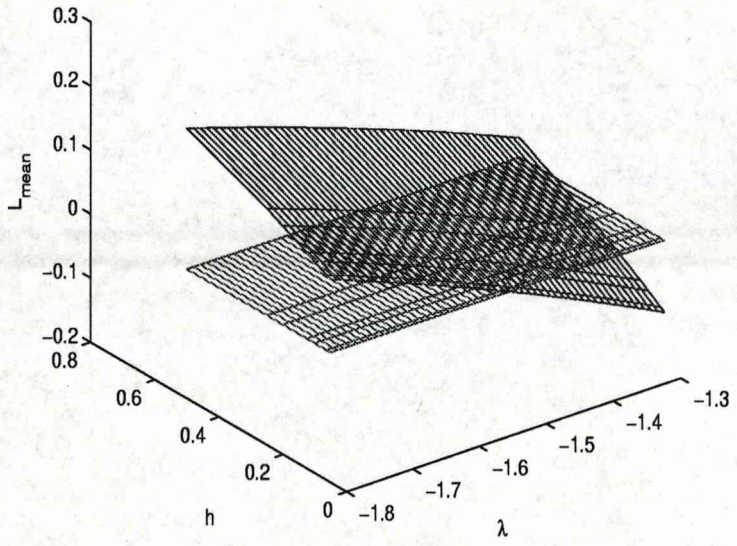


Figure 8.5: L_{mean} against λ and h for $\mu = 0.1$, with the plane $L_{mean} = 0$

If we substitute the values of the coefficients of our equation for $\vartheta = 0$ and $\mu = 0.10$ we find

$$\lambda = -1.570419 - 0.838713h - 0.551683h^2 + \ldots \tag{8.5}$$

Substituting in equation (8.3) shows that this expansion is valid for $-0.737 \leq h \leq 3.268$, a range which clearly includes all of our values of h.

We have repeated this analysis for all 21 of the cases tabled above, and the equations are shown in table 8.4.

In line with the deterministic equation, $\vartheta = 0, 1$ give formulae for λ_{bif} which are a close $\mathcal{O}(h)$ approximation to $-\frac{\pi}{2}$. For $\vartheta = 0.5$, which corresponds to the second order trapezium method, the h coefficients are very small, so we have (to working accuracy) an $\mathcal{O}(h^2)$ approximation to $-\frac{\pi}{2}$. In this case it is also evident that as μ, the noise coefficient, increases the h coefficient in the formula for λ_{bif} becomes more significant.

We also note that by symmetry, we should expect equation (4.3) will clearly give us a similar family of solutions to equation

$$\begin{aligned} dY(t) &= \lambda Y(t-1)dt - \mu Y(t)dW(t), \qquad t \geq 0 \\ Y(t) &= t + \frac{1}{2}, \qquad t \in [-1, 0]. \end{aligned} \tag{8.6}$$

ϑ	μ	Equation
0	0.00	λ = -1.571912 + 0.834695h - 0.547274h^2
	0.05	λ = -1.571528 + 0.835740h - 0.548515h^2
	0.10	λ = -1.570419 + 0.838713h - 0.551683h^2
	0.15	λ = -1.548589 + 0.843807h - 0.557288h^2
	0.20	λ = -1.566100 + 0.851446h - 0.566000h^2
	0.25	λ = -1.562969 + 0.861674h - 0.577721h^2
	0.30	λ = -1.559247 + 0.874770h - 0.593051h^2
0.5	0.00	λ = -1.571133 - 0.004270h - 0.298433h^2
	0.05	λ = -1.570780 - 0.003103h - 0.299150h^2
	0.10	λ = -1.569733 + 0.000278h - 0.300944h^2
	0.15	λ = -1.568048 + 0.006121h - 0.304378h^2
	0.20	λ = -1.565737 + 0.014742h - 0.309660h^2
	0.25	λ = -1.562846 + 0.026218h - 0.316905h^2
	0.30	λ = -1.559441 + 0.041722h - 0.327114h^2
1	0.00	λ = -1.570182 - 0.662504h - 0.733830h^2
	0.05	λ = -1.569866 - 0.661547h - 0.732357h^2
	0.10	λ = -1.568979 - 0.658357h - 0.728407h^2
	0.15	λ = -1.567504 - 0.653117h - 0.721807h^2
	0.20	λ = -1.565492 - 0.645613h - 0.712586h^2
	0.25	λ = -1.562970 - 0.635669h - 0.700930h^2
	0.30	λ = -1.559988 - 0.622769h - 0.687417h^2

Table 8.4: Equation for λ in terms of h at $L = 0$

This indicates that the coefficients in our formulae are very likely to depend only on even powers of μ. Hence, for each ϑ, it makes sense to plot the graphs of each quadratic equation coefficient from table 8.4 against μ^2. Figures 8.6, 8.7 and 8.8 show these graphs for $\vartheta = 1$. The apparent straight lines confirm our conjecture that the quadratic coefficients depend upon even powers of μ. We can calculate the regression formulae for these three coefficients and repeat for the other two ϑ values. In all nine cases $R = 1.000$, giving near perfect linear fits, and confirming the dependency on μ^2. Using linear regression on these figures we obtain

$$\begin{aligned} h^2 \text{ coefficient} &= -0.73357 + 0.51702\mu^2 \\ h \text{ coefficient} &= -0.66280 + 0.43965\mu^2 \\ \text{constant term} &= -1.57183 + 0.14092\mu^2 \\ &\approx -\frac{\pi}{2} + 0.14092\mu^2 \end{aligned}$$

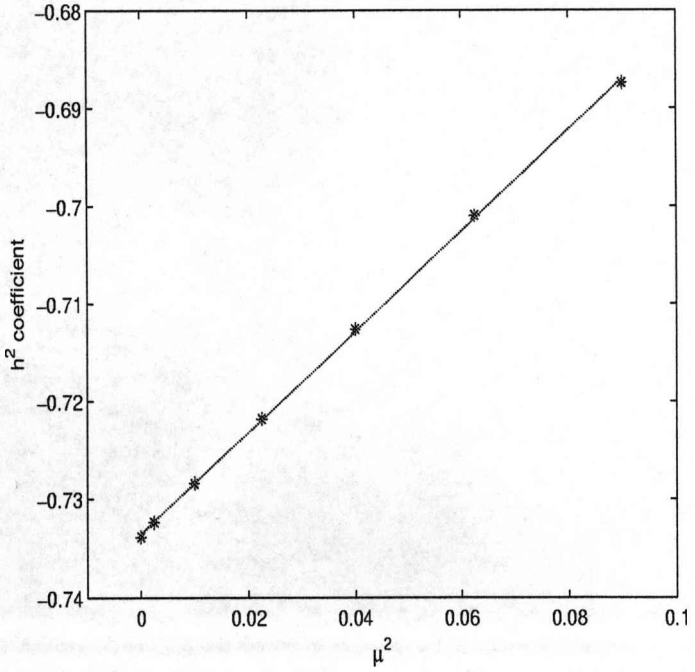


Figure 8.6: Regression line for the h^2 quadratic coefficients against μ^2 for $\vartheta = 1$

We can now obtain approximations to the value of the parameter λ_{bif} where the bifurcations of the linear stochastic delay differential equation occur in terms

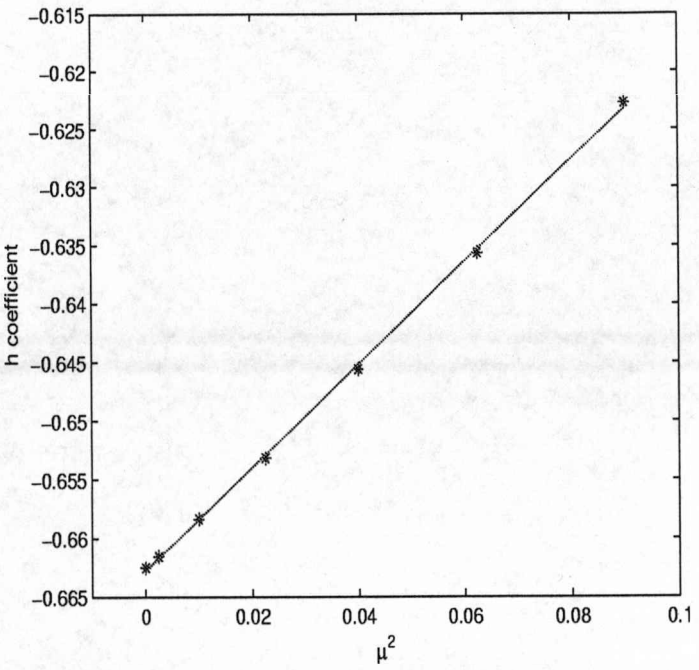


Figure 8.7: Regression line for the h quadratic coefficients against μ^2 for $\vartheta = 1$

of the stepsize h and noise coefficient μ. For $\vartheta = 0$,
(8.7)
$$\lambda = (-1.57183 + 0.14092\mu^2) + (0.83428 + 0.44343\mu^2)h + (-0.54665 - 0.50542\mu^2)h^2$$

For $\vartheta = 0.5$,
(8.8)
$$\lambda = (-1.57105 + 0.13017\mu^2) + (-0.00481 + 0.50704\mu^2)h + (-0.29785 - 0.31503\mu^2)h^2$$

For $\vartheta = 1$,
(8.9)
$$\lambda = (-1.57011 + 0.11346\mu^2) + (-0.66280 + 0.43965\mu^2)h + (-0.73357 + 0.51702\mu^2)h^2$$

These results seem very satisfactory. First we note that, by putting $\mu = 0$, we recover an excellent representation of the known behaviour of these schemes for the deterministic problem. Secondly, we can observe the way in which the presence of noise influences the approximation of the bifurcation point in each of the methods.

For the cases $\vartheta = 0, 1$ the deterministic problem leads to an $\mathcal{O}(h)$ approximation of the exact bifurcation value. We can see that the presence of noise leads to a change in each of the three coefficients in equations (8.7) and (8.9). This means that, in the limit as $h \to 0$ we would expect to obtain an approximation

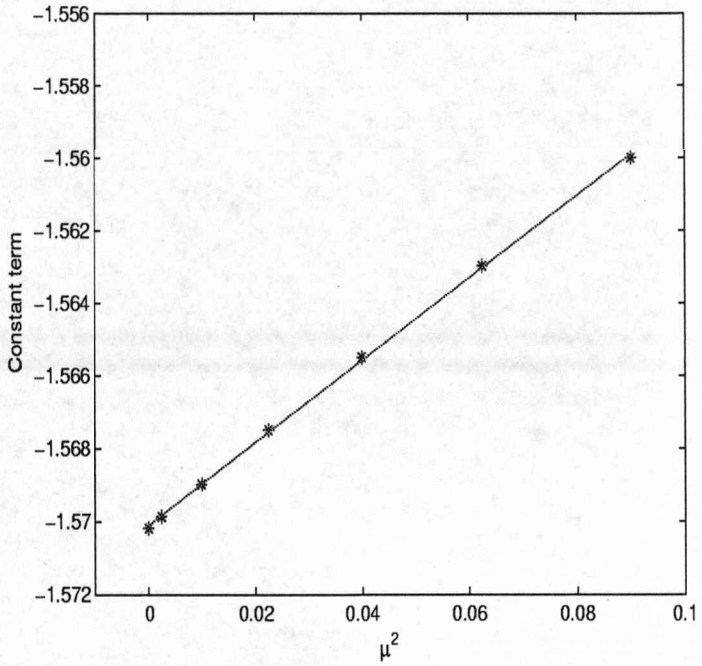


Figure 8.8: Regression lines for the constant quadratic coefficients against μ^2 for $\vartheta = 1$

for the bifurcation value that differs from $-\frac{\pi}{2}$ by an amount proportional to μ^2. During the limiting process, we expect to observe $\mathcal{O}(h)$ convergence.

Now, for the case $\vartheta = 0.5$ one needs to interpret equation (8.8) particularly carefully. If μ is small, then equation (8.8) will provide an *apparent* $\mathcal{O}(h^2)$ rate of convergence as $h \to 0$ in experimental data. It is only when the value of μ is larger that the *true* convergence rate $\mathcal{O}(h)$ will become apparent. This explains why some experiments involving equations with small noise can predict an $\mathcal{O}(h^2)$ approximation to λ. In fact,if $\mu = 0.0974$ then the h term disappears and the method is clearly $\mathcal{O}(h^2)$.

8.4 Conclusions

The results of this chapter provide a systematic approach to analysing the approximate bifurcation values for equation (4.3) and show how the approximations are influenced both by the choice of numerical scheme and its step length and by the level of noise in the equation. There are some observations and questions that are significant and motivate further investigation:

1. The estimates of the bifurcation value (obtained by putting $h = 0$ in equa-

tions (8.7), (8.8) and (8.9)) all indicate that the presence of the noise has induced a change in the bifurcation value.

2. The presence of the μ^2 term in the coefficient of h in (8.8) means that the observed behaviour of approximations might change in a significant way when the level of noise varies. One needs to be particularly careful in applying *small noise* insights to general problems.

Chapter 9

Conclusions and possibilities for further work

In this thesis we have investigated bifurcations in delay differential equations. We have limited our investigations to the linear delay differential equation and stochastic linear delay differential equation with multiplicative instantaneous noise and, to a lesser extent, to the logistic forms. We have investigated the values of the equation parameters at the bifurcation point and in particular how numerical methods affect these parameter values. For the deterministic equations we found that it was possible to approach the problem phenomenologically, by observing the graphs of numerical solutions as we varied our parameters. For the stochastic equations we needed a more dynamical approach and introduced the Lyapunov exponent

$$\Lambda = \lim_{t\to\infty} \sup E(\frac{1}{t}\log|Y(t)|)$$

which we estimated with $L = \frac{\log(S)}{T}$, where $S = \sup_{[T-\epsilon,T]}(|Y|)$. We used L to investigate the bifurcation points.

9.1 Conclusions

1. For the deterministic linear delay differential equation

$$\begin{aligned} dY(t) &= \lambda Y(t-1)dt, \qquad t \geq 0 \\ Y(t) &= t + \frac{1}{2}, \qquad t \in [-1, 0] \end{aligned}$$

it is known that there is a bifurcation for $\lambda = -\frac{\pi}{2}$. We have confirmed that for numerical solutions the bifurcation occurs for $\lambda_{bif} = -\frac{\pi}{2} + \mathcal{O}(h^n)$ where n is the order of the numerical method. The phenomenological approach gave us very precise values.

2. The logistic delay differential equation

$$\begin{aligned} dY(t) &= \lambda Y(t-1)[1+Y(t)]dt, \qquad t \geq 0 \\ Y(t) &= t + \frac{1}{2}, \qquad t \in [-1, 0]. \end{aligned}$$

has a Hopf bifurcation at $\lambda = -\frac{\pi}{2}$. Our phenomenological approach again showed that $\lambda_{bif} = -\frac{\pi}{2} + \mathcal{O}(h^n)$, confirming the work of Wulf [81]. However, the phenomenological values are not as precise as for the linear equation.

3. The phenomenological approach does not give a clear indication of bifurcation values for stochastic delay differential equations. Use of the Lyapunov exponent was far more productive. However, for the logistic equation

$$\begin{aligned} dY(t) &= \lambda Y(t-1)[1+Y(t)]dt + \mu Y(t)dW(t), \qquad t \geq 0 \\ Y(t) &= t + \frac{1}{2}, \qquad t \in [-1, 0]. \end{aligned}$$

we found that where the values of L are positive the values become less predictable. At this point of the investigation we restricted our investigation to the linear equation

$$\begin{aligned} dY(t) &= \lambda Y(t-1)dt + \mu Y(t)dW(t), \qquad t \geq 0 \\ Y(t) &= t + \frac{1}{2}, \qquad t \in [-1, 0]. \end{aligned}$$

4. For the linear SDDE we have shown that, for fixed noise coefficient μ, our approximation to the Lyapunov exponent L varied quadratically with λ. When we solved the quadratic equation $L = 0$ we get an estimate of λ_{bif} which, when $\mu = 0$, corresponds favourably with the phenomenological value for the deterministic equation.

5. For the linear SDDE we have shown that, for fixed noise coefficient μ, our approximation to the Lyapunov exponent L varied quadratically with h.

6. For the linear SDDE we have shown that, for fixed noise coefficient μ, our approximation to the Lyapunov exponent L varied quadratically with both λ and h together.

7. Using the relationship above we have shown that solving $L = 0$ for λ in terms of h gives us a formula for the bifurcation value of λ in the form $\lambda_{bif} \approx -\frac{\pi}{2} + \mathcal{O}(h^n)$, where n is the order of the numerical method.

8. Finally, we have shown that if the form above is $\lambda \approx a + bh + ch^2$ then

$$\begin{aligned} a &\approx a_0 + a_2\mu^2 \\ b &\approx b_0 + b_2\mu^2 \\ c &\approx c_0 + c_2\mu^2 \end{aligned}$$

9.2 Further investigations

This thesis leads to the following questions and possible further investigations.

1. The estimates of the bifurcation value (obtained by putting $h = 0$ in equations (8.7), (8.8) and (8.9)) all indicate that the presence of the noise has induced a change in the bifurcation value. Can this change be established analytically for the underlying SDDE, or is it nevertheless an artefact induced by the numerical scheme?

2. Can a formula be developed that combines equations (8.7), (8.8) and (8.9)) into a single expression with ϑ as parameter? Can such an expression lead to establishing some critical value of ϑ (other than 0, 0.5 and 1) for which the numerical approach displays enhanced properties?

3. How does the choice of the initial function $Y_0(t)$ affect the investigations? (Note that some initial work has been completed on this question and the results are contained in Appendix A

Appendix A

Varying the initial function

We present here some results from repeating our investigations with different initial functions. We will give the results alongside the results from using the initial function applied throughout this thesis for an immediate comparison. We will consider

$$dY(t) = \lambda Y(t-1)dt + \mu Y(t)dW(t), \qquad t \geq 0$$

with our original initial function

$$Y(t) = t + \frac{1}{2}, \qquad t \in [-1, 0]$$

and two further initial functions

$$Y_1(t) = 2000e^{-t+1} - 4000, \qquad t \in [-1, 0]$$

and

$$Y_2(t) = 2000\cos(2t) - 1000, \qquad t \in [-1, 0].$$

We have added a monotonic exponential function and a sinusoidal function, both with large values at $t = 0$, to our simple linear function to ensure we have a much different situation. As this is only a preliminary investigation to verify that the choice of initial function has a minor effect and the choice is not vital to the work of this thesis we have restricted the experiments to the case $\vartheta = 0$ and $\mu = 0.15$, and we have applied all seven step lengths. Figure A.1 shows the graphs of L_{mean} against λ with the seven step lengths for each of our initial functions. The three graphs in figure A.1 look identical. We can now draw the three dimensional graphs of L_{mean} against λ and h for each of our initial functions, and these are shown in figure A.2. Once again we can see no visual difference between the three dimensional graphs for the three initial functions. To find a difference between the diagrams we need to return to our analysis using SPSS. We can find the least squares quadratic regression equations and these are given in table A.1. We again see that $R = 0.999$, showing excellent fits. We can see from table A.1

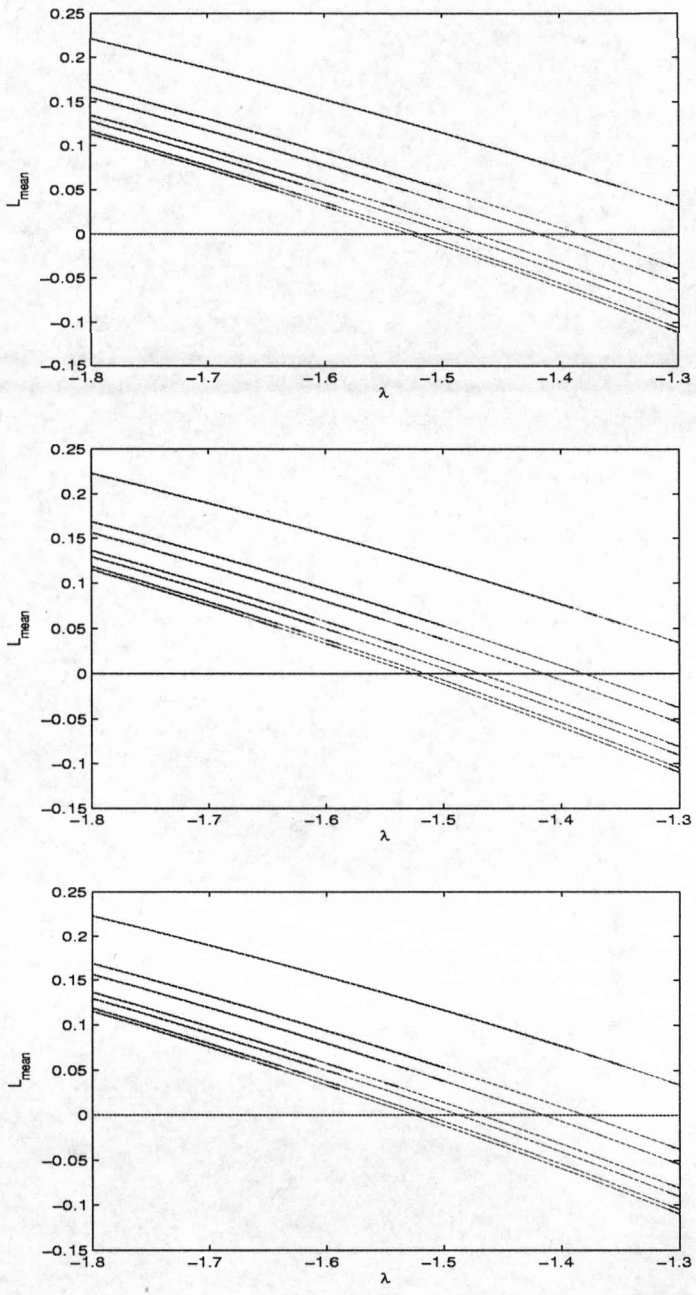


Figure A.1: Mean L against λ for our step lengths with $\vartheta = 0$ and $\mu = 0.15$, for the initial functions:
Top: $Y(t) = t + \frac{1}{2}$
Middle: $Y_1(t) = 2000e^{-t+1} - 4000$
Bottom: $Y_2(t) = 2000\cos(2t) - 1000$

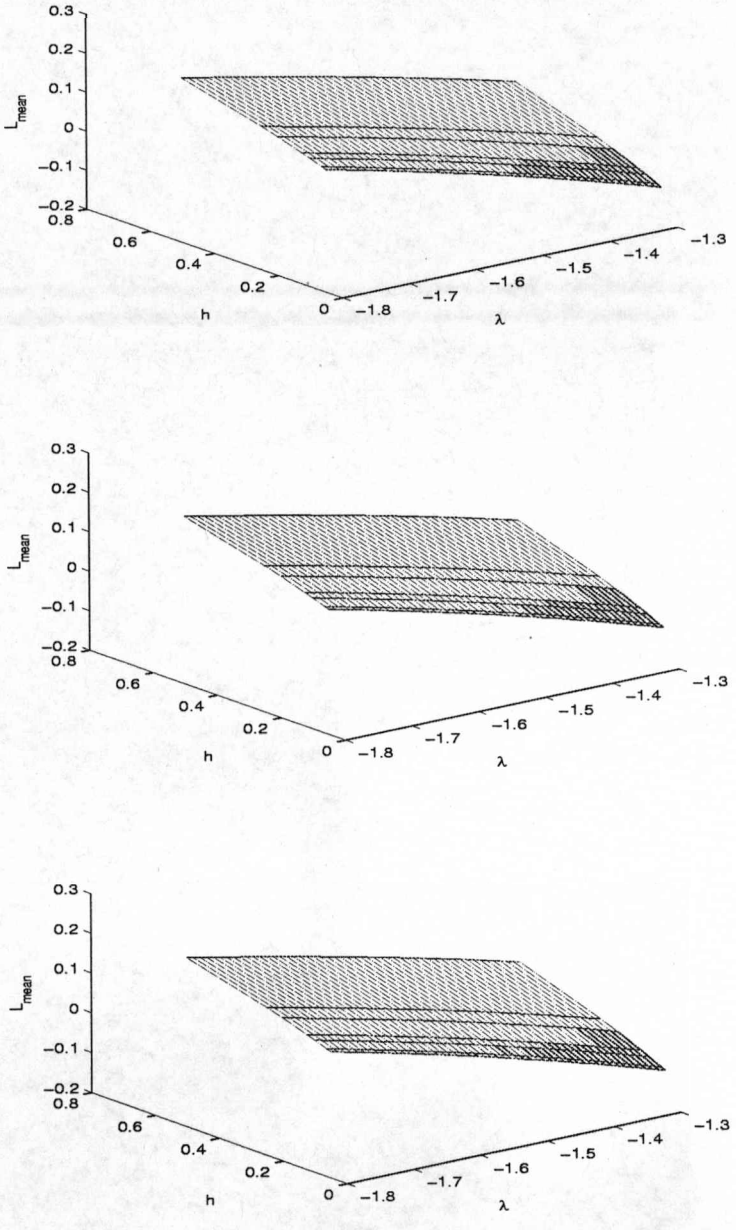


Figure A.2: Mean L against λ and h with $\vartheta = 0$ and $\mu = 0.15$, for the initial functions:
Top: $Y(t) = t + \frac{1}{2}$
Middle: $Y_1(t) = 2000e^{-t+1} - 4000$
Bottom: $Y_2(t) = 2000\cos(2t) - 1000$

Equation	R
$L_{mean} = -.131175\lambda^2 - .141014h^2 - .832148\lambda + .354931h - .982547$	.999
$L_{mean} = -.131225\lambda^2 - .141016h^2 - .832355\lambda + .355011h - .980905$	.999
$L_{mean} = -.131165\lambda^2 - .141025h^2 - .832131\lambda + .354989h - .980992$	.999

Table A.1: Regression formulae for $\vartheta = 0$, $\mu = 0.15$, for the initial functions:
Top: $Y(t) = t + \frac{1}{2}$
Middle: $Y_1(t) = 2000e^{-t+1} - 4000$
Bottom: $Y_2(t) = 2000\cos(2t) - 1000$

Equation
λ = -1.548589 + 0.843807h - 0.557288h^2
λ = -1.564907 + 0.841982h - 0.555042h^2
λ = -1.564215 + 0.841604h - 0.554642h^2

Table A.2: Equation for λ in terms of h at $L_{mean} = 0$ for $\vartheta = 0$, $\mu = 0.15$, for the initial functions:
Top: $Y(t) = t + \frac{1}{2}$
Middle: $Y_1(t) = 2000e^{-t+1} - 4000$
Bottom: $Y_2(t) = 2000\cos(2t) - 1000$

that the change in initial function has caused very small changes in the regression coefficients. We conclude this investigation by solving the equation $L_{mean} = 0$ for λ in terms of h, and show these relationships in table A.2. Once again we can see from table A.2 that the change in initial function has caused very small changes in the coefficients. Finally, we can use the formulae in table A.2 to predict values for λ_{bif} where $\vartheta = 0$ and $\mu = 0.15$ for a particular step length. For instance, if we substitute $h = 0.1$ into the formulae we get the values shown in table A.3. The three initial functions produce values for λ_{bif} which agree to two decimal places. The experiments completed in this appendix confirm that more work needs to be completed on this problem, but the results are consistent with the previous investigations of this project.

Initial function	λ_{bif}
$Y(t) = t + \frac{1}{2}$	-1.4898
$Y_1(t) = 2000e^{-t+1} - 4000$	-1.4863
$Y_2(t) = 2000\cos(2t) - 1000$	-1.4856

Table A.3: λ_{bif} for $h = 0.1$ with $\vartheta = 0$, $\mu = 0.15$ for our three initial functions.

Appendix B

MATLAB code

In this appendix we give the code for the main programs used in this thesis.

1. Used in section 1.2.8 to show that bifurcations in ODEs are not always duplicated in numerical solutions

```
%Program to compare numerical solutions with true solutions
lambda=input('lambda')
theta=input('theta')
h=input('stepsize')
t=0:0.001:1;
z=exp(lambda.*t);          %exact value
n=1/h;
y=zeros(1,n+1);
y(1)=1;
for i=1:n                  %compute numerical solution
    y(i+1)=(1+(1-theta)*h*lambda)/(1-theta*h*lambda)*y(i);
end
t0=0:h:1;
plot(t,z,t0,y)
```

2. Verification of validity of *randn* in section 3.1.2

```
%program to verify randn is N(0,1)
N=input('vector size')
randn('state',1)
dW=randn(1,N);
g=[-5:0.2:5];
hist(dW,g)
hold on
x=[-4:0.01:4];
y=N/5*exp(-x.^2/2)/sqrt(2*pi);
```

```
plot(x,y)
m=mean(dW)
s=std(dW)
```

3. Twenty Brownian paths, with mean path in section 3.1.3

```
randn('state',1);
h=0.01;
T=10;
t=[0:h:T];
n=20;
dW=sqrt(h)*randn(n,T/h);
W=[zeros(n,1) cumsum(dW,2)];
for i=1:n;
    plot(t,W(i,:))
    hold on
end
Wmean=mean(W);
plot(t,Wmean,'r')
```

4. Stochastic integrals of $\int_0^1 tdW$ in section 3.3

```
% Integrals of tdW from 0 to 1
randn('state',1);
step=input('stepsize');
dt=0.00005;
N=1/dt;
dB=sqrt(dt)*randn(1,N);
B=[0 cumsum(dB)];
t=0:dt:1;
ito=0;
strato=0;
length=N/step;
for count =1:step;
    dW=B(count*length+1)-B((count-1)*length+1);
    ito=ito+t((count-1)*length+1)*dW;
    strato=strato+t((count-0.5)*length+1)*dW;
end
ito
strato
```

5. Stochastic integrals of $\int_0^1 WdW$ in section 3.3

```
% Integrals of WdW from 0 to 1
step=input('stepsize');
randn('state',1);
dt=0.00005;
N=1/dt;
dB=sqrt(dt)*randn(1,N);
B=[0 cumsum(dB)];
t=0:dt:1;
ito=0;
strato=0;
length=N/step;
for count =1:step;
    dW=B(count*length+1)-B((count-1)*length+1);
    ito=ito+B((count-1)*length+1)*dW;
    strato=strato+B((count-0.5)*length+1)*dW;
end
ito
strato
B(N)
```

6. To produce the tables of values for $\int_0^1 tdW$ in section 3.3

```
% Integrals of tdW from 0 to 1
randn('state',1);
step=input('stepsize');
dt=0.00005;
t=0:dt:1;
N=1/dt;
itolist=zeros(1,10000);
stratolist=zeros(1,10000);
length=N/step;
for i=1:10000
    dB=sqrt(dt)*randn(1,N);
    B=[0 cumsum(dB)];
    ito=0;
    strato=0;
    for count =1:step;
        dW=B(count*length+1)-B((count-1)*length+1);
        ito=ito+t((count-1)*length+1)*dW;
        strato=strato+t((count-0.5)*length+1)*dW;
    end
    itolist(i)=ito;
    stratolist(i)=strato;
```

```
end
mean(itolist)
mean(stratolist)
std(itolist)
std(stratolist)
```

7. To produce the tables of values for $\int_0^1 tdW$ in section 3.3

```
% Integrals of WdW from 0 to 1
step=input('stepsize');
randn('state',1);
itolist=zeros(1,10000);
stratolist=zeros(1,10000);
dt=0.00005;
N=1/dt;
t=0:dt:1;
for i=1:10000
    dB=sqrt(dt)*randn(1,N);
    B=[0 cumsum(dB)];
    ito=0;
    strato=0;
    length=N/step;
    for count =1:step;
        dW=B(count*length+1)-B((count-1)*length+1);
        ito=ito+B((count-1)*length+1)*dW;
        strato=strato+B((count-0.5)*length+1)*dW;
    end
    itolist(i)=ito;
    stratolist(i)=strato;
end
mean(itolist)
mean(stratolist)
std(itolist)
std(stratolist)
```

8. To produce the diagrams in section 4.2 using forward Euler. One line change is needed for the other three methods.

```
k=input('lambda');
h=input('h');
N=1/h;
X=ones(1,200001*N+1);
Y=ones(1,200001*N+1);
```

```
for x=1:N+1
   t=(x-N-1)/N;
   X(x)=t;
   Y(x)=t+1/2;
end
for x=N+2:200001*N+1
   t=(x-N-1)/N;
   X(x)=t;
   Y(x)=Y(x-1)+h*k*Y(x-1-N); %code for forward Euler
end
plot(X,Y,'-')
```

9. To produce the diagrams 4.7 using Euler-Maruyama.

```
% Euler-Maruyama for dY(t)=lambda*Y(t-1)+ mu*Y(t)dW(t)
lambda = input('lambda =');
mu = input('mu =');
h = input('step size =');
randn('state',100)

T = 200000;
N = 1/h;
M = (T+1)*N;
Y=zeros(1,M);

for i = 1:N+1
    Y(i) = (i-1)/N-0.5;
end

dW = sqrt(h)*randn(1,N*T);

for n = N+1:M
    Y(n+1) = Y(n) + h*lambda*Y(n-N) + mu*Y(n)*dW(n-N);
end

pen = Y(140001*N:150001*N);
fin = Y(190001*N:200001*N);
p = max(abs(pen))
f = max(abs(fin))

t = [-1:h:T];
plot(t,Y)
```

10. To test for strong convergence in section 5.3.1.

```
randn('state',100);
q=input('theta');
inc=[1 2 4 5 8 10 16 20 40];
N=4000;
dt=1/N;
y0=1;
p=5000;
yerr=zeros(p,9);
for s=1:p
    dW=sqrt(dt)*randn(1,N);
    W=cumsum(dW);
    yexact=y0*exp(1+2*W(end));
    for count=1:9
        h=inc(count)*dt;
        n=1/h;
        ytemp=y0;
        for i=1:n
            Winc=sum(dW(inc(count)*(i-1)+1:inc(count)*i));
            ytemp=ytemp*(1+(1-q)*3*h+2*Winc)/(1-q*3*h);
        end;
        yerr(s,count)=abs(ytemp-yexact);
    end;
end;
step=inc/(2*N/100);
A=[ones(9,1), log(step)']
b=log(mean(yerr)')
reg=A\b
t=log(step);
x=reg(2).*t+reg(1);
plot(t,b,'*',t,x)
```

11. To test for weak convergence in section 5.3.2 for $\theta = 0$ and $\theta = 1$.

```
randn('state',100);
q=input('theta');
inc=[2 4 8 10 16 32 64 80];
N=1600;
dt=1/N;
y0=1;
p=20000;
yerr=zeros(8,1);
yend=zeros(p,8);
for s=1:p
```

```
    dW=sqrt(dt)*randn(1,N);
    W=cumsum(dW);
    for count=1:8
        h=inc(count)*dt;
        n=1/h;
        ytemp=y0;
        for i=1:n
            Winc=sum(dW(inc(count)*(i-1)+1:inc(count)*i));
            ytemp=ytemp*(1+(1-q)*2*h+0.1*Winc)/(1-q*2*h);
        end;
        yend(s,count)=ytemp;
    end;

end;
exact=exp(2);
ylast=mean(yend)
yerr=abs(ylast-exact);
step=inc/N;
A=[ones(8,1), log(step)']
b=log(yerr)'
reg=A\b
t=log(step);
x=reg(2).*t+reg(1);
plot(t,b,'*',t,x)
```

12. To test for weak convergence in section 5.3.2 for $\theta = 0.5$

```
randn('state',100);
q=input('theta');
inc=[1 2 4 8 10 16 32 64 80 100 160 200 400];
N=1600;
dt=1/N;
y0=1;
p=400000;
yerr=zeros(13,1);
yend=zeros(p,13);
for s=1:p
    dW=sqrt(dt)*randn(1,N);
    W=cumsum(dW);
    for count=1:13
        h=inc(count)*dt;
        n=1/h;
        ytemp=y0;
```

```
        for i=1:n
            Winc=sum(dW(inc(count)*(i-1)+1:inc(count)*i));
            ytemp=ytemp*(1+(1-q)*2*h+0.1*Winc)/(1-q*2*h);
        end;
        yend(s,count)=ytemp;
    end;

end;
exact=exp(2);
ylast=mean(yend)
yerr=abs(ylast-exact);
step=inc/N
step1=inc(8:13)/N;
A=[ones(6,1), log(step1)'];
b=log(yerr)'
b1=log(yerr(8:13))'
reg=A\b1
t=log(step);
t1=log(step1);
x=reg(2).*t1+reg(1);
%plot(t,b,'*',t1,x)
plot(t,b,'*')
hold on
plot(t1,x)
```

13. To produce the 147 files, each containing the 51 values of L.

```
%EM  Euler-Maruyama method on linear SDDE
% SDDE is  dY = lambda*Y(t-1) dt + mu*Y(t)dW,
% with  Y(t)=t+1/2   -1<=t<=0

theta = input('theta = ?');
mu = input('mu = ?');
Dt = input('step size =?');
T = 5000;  N = 1/Dt;
L = T*N;
for i = 1:N+1
    Ym(i) = (i-1)/N-0.5;
end
iters = 51;
me = zeros(1,iters); mi = zeros(1,iters);
ma = zeros(1,iters); st = zeros(1,iters);
for k = 1:iters
```

```
    lambda(k) = -1.8+ (k-1)*0.01;
    randn('state',100)
    lypend = zeros(1,500);
    for j = 1:500
        p = 0;
        dW = sqrt(Dt)*randn(1,T*N);
        W = cumsum(dW);
        Yi = zeros(1,L);
        Yem = [Ym Yi];
        for n = 1:L-50
        Winc = sum(dW((n-1)+1:n));
        Yem(n+1+N) = Yem(n+N) + (1-theta)*Dt*lambda(k)*Yem(n)
+ theta*Dt*lambda(k)*Yem(n+1) + (mu*Yem(n+N)+nu*Yem(n))*Winc;
        x1 = abs(Yem(N+n+1));
        x2=round(log10(x1));
        Yem(n:N+1+n) = Yem(n:N+1+n).*10^(-x2);
        p = p + x2;
        end
        for n = L-49:L
        Winc = sum(dW((n-1)+1:n));
        Yem(n+1+N) = Yem(n+N) + (1-theta)*Dt*lambda(k)*Yem(n)
+ theta*Dt*lambda(k)*Yem(n+1) + (mu*Yem(n+N)+nu*Yem(n))*Winc;
        x1 = abs(Yem(N+n+1));
        x2=round(log10(x1));
        Yem(L-49:N+1+n) = Yem(L-49:N+1+n).*10^(-x2);
        p = p + x2;
        end
        lend = (log(max(abs(Yem(:,N+1+L-50:N+1+L))))
+p*log(10))/T;
        lypend(j) = lend;
    end
    %hist(lypend,20)
    %plot([-1:Dt:T],Yem), hold off
    me(1,k) = mean(lypend);
    mi(1,k) = min(lypend);
    ma(1,k) = max(lypend);
    st(1,k) = std(lypend);
end
m1=mi';
m2=me';
m3=ma';
m4=st';
matlin=[m1 m2 m3 m4]
```

```
z=zeros(1,iters);
plot(lambda,me,lambda,mi,lambda,ma,lambda,z)
```

14. To plot the 3 dimensional surface L against λ and h in section 8.2.

```
lambda=-1.8:0.01:-1.3;
h=[.5 .25 .2 .125 0.1 0.0625 0.05];
load L131.dat
for i = 1:7
    [X,Y]=meshgrid(lambda,h(i));
    plot3(X,Y,L131([1+51*(i-1):51*i],4))
    hold on
end
```

followed by

```
t = input('type, min=3,mean=4,max=5,st dev=6');
lambda=-1.8:0.01:-1.3;
h=[.5 .25 .2 .125 .1 .0625 .05];
load L131.dat
for i = 1:51
    [X,Y]=meshgrid(lambda(i),h);
    plot3(X,Y,L131([i:51:i+306],t))
    hold on
end
```

Bibliography

B.1 Thesis bibliography

[1] Arnold, L., *Random Dynamical Systems*, Springer, Berlin (2000).

[2] L. Arnold & V. Wihstutz (eds), *Lyapunov Exponents*, Springer (1986).

[3] C.T.H. Baker, G.A.Bocharov, J.M. Ford, P.M. Lumb, S.J. Norton, C.A.H. Paul, T. Junt, P. Krebs, B. Ludewig *Computational approaches to parameter estimation and model selection in immunology*, J.Comput. Appl. Math.,**184**, (2005), 50–76.

[4] C.T.H. Baker, G.A.Bocharov & F.A. Rihan, *A Report on the use of Delay Differential Equations in Numerical Modelling in the Biosciences*, University of Manchester, Numerical Analysis Report No 343, http://www.ma.man.ac.uk/mccm/mccm.html

[5] C.T.H. Baker, J.M. Ford & N.J. Ford, *Bifurcations in approximate solutions of stochastic delay differential equations*, Int. J. Bifurcation and Chaos, **14** (2004) 2099–3021.

[6] C.T.H. Baker, E. Buckwar, *Numerical analysis of explicit one-step methods for stochastic delay differential equations*, LMS J. Comput. Math., **3** (2000) 315–335.

[7] C.T.H. Baker, E. Buckwar, *Continuous Θ-methods for the stochastic pantograph equation*, Electron. Trans. Numer. Anal., **11** (2000), 131–151.

[8] C.T.H. Baker, E. Buckwar, *Exponential stability in p-th mean of solutions, and of convergent Euler type solutions, of stochastic delay differential equations*, J.Comput. Appl. Math.,**184**, (2005), 404–427.

[9] Baker, C.T.H. & Buckwar, E., *Introduction to the Numerical Analysis of Stochastic Delay Differential equations*, University of Manchester, Numerical Analysis Report No 345, http://www.ma.man.ac.uk/mccm/mccm.html

[10] C.T.H. Baker, C.A.H. Paul C.A.H., D.R/ Willé, *Issues in the numerical solution of evolutionary delay differential equations*, Adv. Comput. Math., **3** (1995), 171–196.

[11] R. Bellman, K.L.Cooke, *Differential-Difference Equations*, Academic Press, 1963.

[12] A. Bellen, M. Zennaro, *Numerical methods for delay differential equations*, Oxford Science Publications, 2003.

[13] Beretta E., Kolmanovskii V. & Shaiket, L. *Stability of epidemic modelling with time delays influenced by stochastic perturbations*, Mathematics and Computers in Simulation, 45 (1998) 269–277.

[14] Braun, M. *Differential Equations and their Applications*, Springer-Verlag, New York (1975).

[15] Brugnano L. & Trigiante D., *Solving Differential Problems by Multistep Initial and Boundary Value Methods*, Gordon and Breach Science Publishers, Australia (1998).

[16] Burghes, D.N. & Borrie, M.S. *Modelling with Differential Equations*, Ellis Horwood, Chichester (1981).

[17] K. Burrage, P.M. Burrage, T. Tian, *Numerical methods for strong solutions of SDEs*, Proc. R. Society Lond. Ser. A **460** (2004) 373–402.

[18] Butcher, J.C. *Numerical Methods for Ordinary Differential Equations*, Wiley, Chichester (2003).

[19] Carletti, M., Burrage, K. & Burrage, P. M., *Numerical Simulation of Stochastic Ordinary Differential Equations in Biomathematical Modelling*, Mathematics and Computers in Simulation 64 (2004),271–277

[20] Carletti, M., *Numerical Solutions of Stochastic Differential Problems in the Biosciences*, Journal of Computational Methods in Science and Engineering, Vol 2, No 3, 2 (2003), 1–13

[21] V. Chinnathambi & S. Rajasekar, *Statistical dynamics at critical bifurcations in Duffing-van der Pol oscillator*, Paramana- Journal of Physics, **52** (1999) 561–577.

[22] Coddington, E.A. *An Introduction to Ordinary Differential Equations*, Dover Publishers, New York (1961).

[23] C. Corduneanu, *Principles of Differential and Integral Equations*, AMS Chelsea Publishing, 1977.

[24] Diekmann, O., van Gils, S.A., Verduyn Lunel, S.M., & Walther, H-O., *Delay Equations: functional-, complex- and nonlinear analysis*, Springer Verlag, New York, 1995.

[25] E.J. Doedel, H.B. Keller, J.P. Kernevez, *Numerical analysis and control of bifurcation problems (I) bifurcation in finite dimensions*, International Journal of Bifurcation and Chaos, **1** (1991) 493–520.

[26] E.J. Doedel, H.B. Keller, J.P. Kernevez, *Numerical analysis and control of bifurcation problems (II) bifurcation in infinite dimensions*, International Journal of Bifurcation and Chaos, **1** (1991) 745–772.

[27] Doornik, J.A., *An improved Ziggurat method to generate normal random samples*, University of Oxford.

[28] Driver, R.D. *Ordinary and Applied Differential Equations*, Springer-Verlag, New York (1997).

[29] Edwards, C.H. & Penney, D.E. *Differential Equations and Boundary Value Problems*, Prentice Hall, New Jersey (1996).

[30] K. Engelborghs, *Numerical bifurcation analysis of delay differential equations* Thesis. Katholieke Universiteit Leuven, Belgium, 2000.

[31] K. Engelborghs, D. Roose, *Numerical computation of stability and detection of Hopf bifurcations of steady state solutions of delay differential equations,* Adv. Comput. Math., **10** (1999), 271–289.

[32] K. Engelborghs, T. Luzyanina, D. Roose, *Numerical bifurcation analysis of delay differential equations*, J. Comput. Appl. Math., **125**, 2000, 265–275.

[33] Eriksson, K., Estep, D., Hansbo, P. & Johnson, C., *Computational Differential Equations*, Cambridge University Press, Cambridge (1996).

[34] N.J. Ford & S.J. Norton, *Numerical investigation of D-bifurcations for a Stochastic Delay Logistic Equation*, Stochastic Dynamics, Vol. 5, No. 2 (2005) 211-222.

[35] N.J. Ford & S.J. Norton, *Using approximations to Lyapunov exponents to predict changes in dynamical behaviour in numerical solutions to stochastic delay differential equations.*, Algorithm for Approximation, A.Aske and J.Levesley (eds), Springer-Verlag, Heidelberg, 2006, 311-320.

[36] N.J. Ford & S.J. Norton, *Noise-induced changes to the behaviour of semi-implicit Euler methods for stochastic delay equations undergoing bifurcation.*, J. Comput. Appl. Math., (2008), doi:10.1016/jcam.2008.04.017.

[37] Ford, N.J. & Wulf, V., *The use of boundary locus plots in the identification of bifurcation points in the numerical approximation of delay differential equations*, J. Comput. Appl. Math., **111** (1999) 153-162.

[38] Ford, N.J. & Wulf, V., *Hopf bifurcation for numerical approximations to the delay logistic equation.* Internat. J. Appl. Sci. Comput., **6** (1999), 167–172.

[39] Ford, N.J. & Wulf, V., *Numerical Hopf bifurcation for a class of delay differential equations*, J. Comput. Appl. Math., **115** (2000), 601-616.

[40] Ford, N.J. & Wulf, V., *How do numerical methods perform for delay differential equations undergoing a Hopf bifurcation?* J. Comput. Appl. Math., **125** (2000), 277-285.

[41] Fulford, G., Forrester, P. & Jones, A. *Modelling with Differential and Difference Equations*, Cambridge University Press, Cambridge (1997).

[42] Giles, M., *Monte Carlo Evaluation of Sensitivities in Computational Finance*, Report No 07/12, Oxford University Computer Laboratories.

[43] Gilseng, H. & Shardlow, T., *SDElab: Stochastic Differential Equations with MATLAB*, University of Manchester MIMS Eprint 2006.1, http://www.manchester.ac.uk/mims/eprints.

[44] Golec, G. & Sathananthan, S., *Stability Analysis of a Stochastic Logistic Model*, Mathematics and Computer Modelling 38 (2003), 585–593.

[45] Gopalsamy, K., *Stability and oscillations in delay differential equations of population dynamics*, Kluwer Academic, London, 1992.

[46] W.J.F. Govaerts, *Numerical methods for bifurcations of dynamic equilibria*, SIAM Publications, Philadelphia, 2000.

[47] Hairer E., Norsett, S.P. & Wanner, G., *Solving Ordinary Differential Equations 1*, Springer, Berlin (2000).

[48] Hale, J. & Verduyn Lunel, S.M., *Introduction to Functional Differential Equations*, Springer-Verlag, New York, 1993.

[49] Higham, D.J., *An algorithmic introduction to numerical simulation of stochastic differential equations*, SIAM rev 43, 2001, 525–546.

[50] D.J. Higham, *Mean-square and asymptotic stability of the stochastic theta method*, SIAM J. Numer. Anal., 38, (2000), 753–769.

[51] D.J. Higham, X. Mao, A.M. Stuart, *Strong convergence of numerical methods for nonlinear stochastic differential equations*, SIAM J. Numer. Anal. 40 (2002) 1041–1063.

[52] Higham, D. J., *Option Valuation for Scientific Computing Students*, University of Strathclyde Mathematics Research Report 01 (2004).

[53] Holmes, M.H. *Introduction to Numerical Methods in Differential Equations*, Springer, New York (2007).

[54] Y. Hu, S. E. A. Mohammed, F. Yan, *Discrete-time Approximations of Stochastic Differential Systems with Memory* , Unpublished typescript (electronic, URL: `http://salah.math.siu.edu/`)

[55] Y. Hu, S. E. A. Mohammed, F. Yan, *Discrete-time Approximations of Stochastic Differential Systems: The Milstein Scheme* , Ann. Probab., Vol 32, Number 1A (2004), 265–314.

[56] Iserles A., *A first course in the Numerical Analysis of Differential Equations*, Cambridge University Press, Cambridge (1996).

[57] Klebaner F. C., *Introduction to Stochastic Calculus with Applications*,Imperial College Press, London (1998).

[58] Kloeden, P.E. & Platen E., *Numerical Solution of Stochastic Differential Equations*, Springer, Berlin 1999.

[59] Kolmanovskii, V. & Myshkis, A. *Introduction to the Theory and Applications of Functional Differential Equations*, Kluwer Academic Publishers, Dordrecht (1999).

[60] Lambert, J.D., *Numerical Methods for Ordinary Differential Systems*, Wiley, Chichester (1991).

[61] Liu, M., Wanrong, C. & Zhacheng, F, *Convergence and Stability of the Semi-implicit Euler Method for a Linear Stochastic Differential Delay Equation*, Journal of Computational and Applied Mathematics 170 (2004), 255–268.

[62] Leong, P.H.W., Zhang, G., Lee, D.-U., Luk, W. & Villasenor, J. *A comment on the implementation of the Ziggurat method*, Journal of Statistical Software, **12** (2005) 1-4.

[63] T. Luzyanina, D. Roose, *Numerical stability analysis and computation of Hopf bifurcation points for delay differential equations,* J. Comput. Appl. Math., **72** (1996), 379–392.

[64] Mao, X., *Stochastic Differential Equations and their Applications*, Horwood Publishing, Chichester (1997).

[65] Marsaglia, G. & Tsang, W.W., *The Ziggurat method for generating random variables*, J. Statistical Software, (2000), 5(8).

[66] Mikosch, T., *Elementary Stochastic Calculus with finance in view*, World Scientific Publishing Co., Singapore 1998.

[67] Milstein, G.N., *Numerical Integration of Stochastic Differential Equations*, Kluwer Academic Press, Dordrecht (1995).

[68] S-E.A. Mohammed & M.K.R. Scheutzow, *Lyapunov exponents of linear stochastic functional differential equations driven by semimartingales I*, Ann. Inst, H. Poincarre Probab. Statist. **32** (1996) 69–105.

[69] S-E.A. Mohammed & M.K.R. Scheutzow, *Lyapunov exponents of linear stochastic functional differential equations II*, Ann. Probability **25** (1997) 1210–1240.

[70] Murray, J.D., *Mathematical Biology, 1: An Introduction* (Third Edition) Springer, New York 2002.

[71] Norton S. J. & Ford N. J. *Predicting changes in dynamical behaviour in solutions to stochastic delay differential equations*, Communications on Pure and Applied Analysis, Vol 5, No 2, June 2006, 367–382

[72] Quarteroni, A., Sacco, R. & Saleri, F. *Numerical Mathematics*, Springer, New York (2000).

[73] Oksendal, B., *Stochastic Differential Equations*, (Fifth Edition) Springer, Berlin 2000.

[74] S. Rajasekar & V. Chinnathambi, *Variation of certain statistical quantities near critical bifurcations, Chaos solitons and fractals*, **11** (2000) 859–870.

[75] Reiss, M., *Stochastic Delay Differential Equations*, Lecture notes, Humboldt University, Berlin, 2002/03.

[76] Shampine, L.F. *Numerical Solutions of Ordinary Differential Equations*, Chapman and Hall, New york (1994).

[77] Sheskin, D.J. *Handbook of Parametric and Nonparametric Statistical Procedures (Second Edition)*, Chapman and Hall, Boca Raton (2000).

[78] Steele, J. M., *Stochastic Calculus and Financial Applications*, Springer, New York 2001.

[79] D. Talay, *The Lyapunov exponent of the Euler scheme for stochastic differential equations*, H. Crauel, M. Grundlach (eds), Stochastic Dynamics, Springer (1999) 241–258.

[80] R.C.L. Wolff, *Local Lyapunov exponents- looking closely at chaos*, J. Roy. Stat. Soc., **54** (1992) 353–371.

[81] Wulf, V. *Numerical analysis of delay differential equations undergoing a Hopf bifurcation*, PhD thesis, University of Liverpool, 1999

B.2 Bibliography for software given in [12]

[82] Bocharov, G.A., Marchuk, G.I. & Romanyukha, A.A. *Numerical solution of LMMs of stiff delay-differential system modelling an immune-response*, Numer. Math. 73 (1992), 223–234.

[83] Butcher, J.C. *The Adaption of Stride to delay differential equations*, Appl. Numer. Math. 9 (1992), 415–425.

[84] Corwin, S.P., Sarafyan, D. & Thompson, R. *DKLag6: a code based on continuously imbedded sixth-order Runge-Kutta methods for the solution of state-dependent functional differential equations*, Appl. Numer. Math. 24 (1997), 319–330.

[85] Enright, W.H. & Hayashi, H. *A delay differential equation solver based on a continuous Runge-Kutta method with defect control*, Numer. Algorithms 16 (1997), 349–364.

[86] Guglielmi, N. & Hairer, E. *Implementing Radau IIA methods for stiff delay differential equations*, Computing 67 (2001), 1–12.

[87] Jackiewicz, z. & Lo, E. *The numerical solutionof neutral functional differential equations by Adams predictor-corrector methods*, Appl. Numer. Math. 8 (1991), 477–491.

[88] Neves, K.W. *Automatic integration of functional differential equations: an approach*, ACM Trans. Math. Software 1 (1975), 357–368.

[89] Neves, K.W. *Algorithm 497 - Automatic integration of functional differential equations*, ACM Trans. Math. Software 1 (1975), 369–371.

[90] Shampine, L.F. & Thompson, S. *Solving DDEs in MATLAB*, Appl. Numer. Math. 37 (2001), 441–458.

[91] Tavernini, L. *A user-friendly interactive Turbo Pascal simulation toolkit*, Simulation 53 (1989), 45–55.

[92] Tavernini, L. *Continuous-time Modelling and Simulation: Using Turbo Pascal and CTMS/TP*, Gordon & Breach, Amsterdam, 1996.

[93] Weiner, R. & Strehmel, K. *A type sensitive code for delay differential equations based on adaptive and explicit Runge-Kutta interpolation methods*, Computing 40 (1988), 255–265.

[94] Willé, D.R. & Baker, C.T.H. *DELSOL - a numerical code for the solution of systems of delay differential equations*, Appl. Numer. Math. 9 (1996), 131–148.

B.3 Bibliography for paper [3]

[95] Akaike H., *A new look at the statistical model identification*, IEEE Transactions on Automatic control, 19 (1974) 716–723.

[96] Altman, J.D., Moss, P.A.H., Goulder, P.J.R., Barouch, D.H., McHeyzer-Williams, M.G., Bell, J.I., McMichael, A.J., Davis, M.M. *Phenotypic analysis of antigen-specific T lymphocytes*, Science, 274, (1996), 94–96.

[97] Audoly, S., Belllu, G., D'Angio, L., Saccomni, M., Cobelli, C. *Global identifiablity of nonlinear biological systems*, IEEE Trans. Biomedical Eng., 48, (2001), 55–65.

[98] Baker, C.T.H., Parmuzin, E.I., *Identification of the initial function for delay differential equations: Parts I, II, III*, MCCM Tech. Rep., 431, 443 & 444,ISSN 1360-1725, University of Manchester (2004).

[99] Battegay, M., Cooper, S., Althage,A., Banziger, H., Hengartner, H., Zinkernagel, R.M. *Quantification of lymphocytic choriomeningitis virus with an immunological focus assay in 24- or 96-well plates*, J. Virol. Methods, 33, (1991), 191–198.

[100] Borghans, J.A., Taams, L.S., Wauben, M.H.M., De Boer, R.J., *Competition for antigenic sites during T cell proliferation: a mathematical interpretation of in vitro data*, Proc. Natl. Acad. Sci. USA., 96, (1999) 10782–10787.

[101] Burnet, F.M. *The Clonal Selection Theory of Acquired Immunity*, (Cambridge University Press, 1959).

[102] Burnham, K.P., Anderson, D.R., *Model selection and inference - a practical information-theoretic approach*, Springer, New York, 1998.

[103] Conn, A.R., Gould, N., Toint, P.L., *LANCELOT: a Fortran package for large-scale nonlinear optimization*, Springer Series, Computational Mathematics, vol. 17, Springer Verlag, New York, 1992.

[104] De Boer, R.J., Oprea, M., Antia, R., Murali-Krishna, K., Ahmed, R., Perelson, A.S. *Recruitment times, proliferation, and apoptosis rates during the $CD8^+$ T-cell response to lymphocytic choriomeningitis virus*, J. Virology, 75, (2001). 10663–10669.

[105] Ehl, S., Klenerman, P., Zinkernagel, R.M., Bocharov, G. *The impact of variation in the number of CD8$^+$ T-cell precursors on the outcome of virus infection*, Cellular Immunology, 189, (1998) 67–73.

[106] Hairer, E., Nørsett, S.P., Wanner, G., *Solving ordinary differential equations. I. Nonstiff problems (second edition)*, Springer-Verlag, Berlin, 1993.

[107] Kullback, S., Leibler, R.A. *On information and sufficiency*, Ann. Math. Stat., 22, (1951), 79–86.

[108] Numerical Algorithms Group, *The NAg* FORTRAN *Library* `http://www.nag.co.uk/numeric/Fortran_Libraries.asp`.

[109] Paul, C.A.H., *A User Guide to Archi*, MCCM Rep. 283, University of Manchester.
`http://www.maths.man.ac.uk/~chris/reports/rep283.pdf`

[110] Schwarz, G. *Estimating the dimension of a model*, The Annals of Statistics, 6, (1978), 461–464.

[111] Zinkernagel, R.M., *Lymphocytic Choriomeningitis Virus and Immunology*, Current Topics in Microbiology and Immunology, 263 (2002) 1–5.